The motion of point particles in curved spacetime

Eric Poisson, Adam Pound, and Ian Vega
Department of Physics, University of Guelph, Guelph, Ontario, Canada N1G 2W1

Major update of Living Reviews article, with final revisions (September 25, 2011)

Abstract

This review is concerned with the motion of a point scalar charge, a point electric charge, and a point mass in a specified background spacetime. In each of the three cases the particle produces a field that behaves as outgoing radiation in the wave zone, and therefore removes energy from the particle. In the near zone the field acts on the particle and gives rise to a self-force that prevents the particle from moving on a geodesic of the background spacetime. The self-force contains both conservative and dissipative terms, and the latter are responsible for the radiation reaction. The work done by the self-force matches the energy radiated away by the particle.

The field's action on the particle is difficult to calculate because of its singular nature: the field diverges at the position of the particle. But it is possible to isolate the field's singular part and show that it exerts no force on the particle — its only effect is to contribute to the particle's inertia. What remains after subtraction is a regular field that is fully responsible for the self-force. Because this field satisfies a homogeneous wave equation, it can be thought of as a free field that interacts with the particle; it is this interaction that gives rise to the self-force.

The mathematical tools required to derive the equations of motion of a point scalar charge, a point electric charge, and a point mass in a specified background spacetime are developed here from scratch. The review begins with a discussion of the basic theory of bitensors (part I). It then applies the theory to the construction of convenient coordinate systems to chart a neighbourhood of the particle's word line (part II). It continues with a thorough discussion of Green's functions in curved spacetime (part III). The review presents a detailed derivation of each of the three equations of motion (part IV). Because the notion of a point mass is problematic in general relativity, the review concludes (part V) with an alternative derivation of the equations of motion that applies to a small body of arbitrary internal structure.

Contents

1 Introduction and summary

1.1 Invitation

The motion of a point electric charge in flat spacetime was the subject of active investigation since the early work of Lorentz, Abrahams, Poincaré, and Dirac [1], until Gralla, Harte, and Wald produced a definitive derivation of the equations of motion [2] with all the rigour that one should demand, without recourse to postulates and renormalization procedures. (The field's early history is well related in Ref. [3].) In 1960 DeWitt and Brehme [4] generalized Dirac's result to curved spacetimes, and their calculation was corrected by Hobbs [5] several years later. In 1997 the motion of a point mass in a curved background spacetime was investigated by Mino, Sasaki, and Tanaka [6], who derived an expression for the particle's acceleration (which is not zero unless the particle is a test mass); the same equations of motion were later obtained by Quinn and Wald [7] using an axiomatic approach. The case of a point scalar charge was finally considered by Quinn in 2000 [8], and this led to the realization that the mass of a scalar particle is not necessarily a constant of the motion.

This article reviews the achievements described in the preceding paragraph; it is concerned with the motion of a point scalar charge q, a point electric charge e, and a point mass m in a specified background spacetime with metric $g_{\alpha\beta}$. These particles carry with them fields that behave as outgoing radiation in the wave zone. The radiation removes energy and angular momentum from the particle, which then undergoes a radiation reaction — its world line cannot be simply a geodesic of the background spacetime. The particle's motion is affected by the near-zone field which acts directly on the particle and produces a *self-force*. In curved spacetime the self-force contains a radiation-reaction component that is directly associated with dissipative effects, but it contains also a conservative component that is not associated with energy or angular-momentum transport. The self-force is proportional to q^2 in the case of a scalar charge, proportional to e^2 in the case of an electric charge, and proportional to m^2 in the case of a point mass.

In this review we derive the equations that govern the motion of a point particle in a curved background spacetime. The presentation is entirely self-contained, and all relevant materials are developed *ab initio*. The reader, however, is assumed to have a solid grasp of differential geometry and a deep understanding of general relativity. The reader is also assumed to have unlimited stamina, for the road to the equations of motion is a long one. One must first assimilate the basic theory of bitensors (part I), then apply the theory to construct convenient coordinate systems to chart a neighbourhood of the particle's world line (part II). One must next formulate a theory of Green's functions in curved spacetimes (part III), and finally calculate the scalar, electromagnetic, and gravitational fields near the world line and figure out how they should act on the particle (part IV). A dedicated reader, correctly skeptical that sense can be made of a point mass in general relativity, will also want to work through the last portion of the review (part V), which provides a derivation of the equations of motion for a small, but physically extended, body; this reader will be reassured to find that the extended body follows the same motion as the point mass. The review is very long, but the satisfaction derived, we hope, will be commensurate.

In this introductory section we set the stage and present an impressionistic survey of what the review contains. This should help the reader get oriented and acquainted with some of the ideas and some of the notation. Enjoy!

1.2 Radiation reaction in flat spacetime

Let us first consider the relatively simple and well-understood case of a point electric charge e moving in flat spacetime [3, 9, 10]. The charge produces an electromagnetic vector potential A^α that satisfies the wave equation

$$\Box A^\alpha = -4\pi j^\alpha \tag{1.1}$$

together with the Lorenz gauge condition $\partial_\alpha A^\alpha = 0$. (On page 294 Jackson [9] explains why the term "Lorenz gauge" is preferable to "Lorentz gauge".) The vector j^α is the charge's current density, which is formally written in terms of a four-dimensional Dirac functional supported on the charge's world line: the density is zero everywhere, except at the particle's position where it is infinite. For concreteness we will imagine that the particle moves around a centre (perhaps another charge, which is taken to be fixed) and that it emits outgoing radiation. We expect that the charge will undergo a radiation reaction and that it

will spiral down toward the centre. This effect must be accounted for by the equations of motion, and these must therefore include the action of the charge's own field, which is the only available agent that could be responsible for the radiation reaction. We seek to determine this self-force acting on the particle.

An immediate difficulty presents itself: the vector potential, and also the electromagnetic field tensor, diverge on the particle's world line, because the field of a point charge is necessarily infinite at the charge's position. This behaviour makes it most difficult to decide how the field is supposed to act on the particle.

Difficult but not impossible. To find a way around this problem we note first that *the situation considered here, in which the radiation is propagating outward and the charge is spiraling inward, breaks the time-reversal invariance of Maxwell's theory.* A specific time direction was adopted when, among all possible solutions to the wave equation, we chose $A^\alpha_{\rm ret}$, the *retarded solution*, as the physically relevant solution. Choosing instead the *advanced solution* $A^\alpha_{\rm adv}$ would produce a time-reversed picture in which the radiation is propagating inward and the charge is spiraling outward. Alternatively, choosing the linear superposition

$$A^\alpha_{\rm S} = \frac{1}{2}\bigl(A^\alpha_{\rm ret} + A^\alpha_{\rm adv}\bigr) \tag{1.2}$$

would restore time-reversal invariance: outgoing and incoming radiation would be present in equal amounts, there would be no net loss nor gain of energy by the system, and the charge would undergo no radiation reaction. In Eq. (1.2) the subscript 'S' stands for 'symmetric', as the vector potential depends symmetrically upon future and past.

Our second key observation is that while the potential of Eq. (1.2) does not exert a force on the charged particle, *it is just as singular as the retarded potential in the vicinity of the world line.* This follows from the fact that $A^\alpha_{\rm ret}$, $A^\alpha_{\rm adv}$, and $A^\alpha_{\rm S}$ all satisfy Eq. (1.1), whose source term is infinite on the world line. So while the wave-zone behaviours of these solutions are very different (with the retarded solution describing outgoing waves, the advanced solution describing incoming waves, and the symmetric solution describing standing waves), the three vector potentials share the same singular behaviour near the world line — all three electromagnetic fields are dominated by the particle's Coulomb field and the different asymptotic conditions make no difference close to the particle. This observation gives us an alternative interpretation for the subscript 'S': it stands for 'singular' as well as 'symmetric'.

Because $A^\alpha_{\rm S}$ is just as singular as $A^\alpha_{\rm ret}$, removing it from the retarded solution gives rise to a potential that is well behaved in a neighbourhood of the world line. And because $A^\alpha_{\rm S}$ is known not to affect the motion of the charged particle, *this new potential must be entirely responsible for the radiation reaction.* We therefore introduce the new potential

$$A^\alpha_{\rm R} = A^\alpha_{\rm ret} - A^\alpha_{\rm S} = \frac{1}{2}\bigl(A^\alpha_{\rm ret} - A^\alpha_{\rm adv}\bigr) \tag{1.3}$$

and postulate that it, and it alone, exerts a force on the particle. The subscript 'R' stands for 'regular', because $A^\alpha_{\rm R}$ is nonsingular on the world line. This property can be directly inferred from the fact that the regular potential satisfies the homogeneous version of Eq. (1.1), $\Box A^\alpha_{\rm R} = 0$; there is no singular source to produce a singular behaviour on the world line. Since $A^\alpha_{\rm R}$ satisfies the homogeneous wave equation, it can be thought of as a free radiation field, and the subscript 'R' could also stand for 'radiative'.

The self-action of the charge's own field is now clarified: a singular potential $A^\alpha_{\rm S}$ can be removed from the retarded potential and shown not to affect the motion of the particle. What remains is a well-behaved potential $A^\alpha_{\rm R}$ that must be solely responsible for the radiation reaction. From the regular potential we form an electromagnetic field tensor $F^{\rm R}_{\alpha\beta} = \partial_\alpha A^{\rm R}_\beta - \partial_\beta A^{\rm R}_\alpha$ and we take the particle's equations of motion to be

$$ma_\mu = f^{\rm ext}_\mu + eF^{\rm R}_{\mu\nu}u^\nu, \tag{1.4}$$

where $u^\mu = dz^\mu/d\tau$ is the charge's four-velocity [$z^\mu(\tau)$ gives the description of the world line and τ is proper time], $a^\mu = du^\mu/d\tau$ its acceleration, m its (renormalized) mass, and $f^\mu_{\rm ext}$ an external force also acting on the particle. Calculation of the regular field yields the more concrete expression

$$ma^\mu = f^\mu_{\rm ext} + \frac{2e^2}{3m}\bigl(\delta^\mu_{\ \nu} + u^\mu u_\nu\bigr)\frac{df^\nu_{\rm ext}}{d\tau}, \tag{1.5}$$

in which the second term is the self-force that is responsible for the radiation reaction. We observe that the self-force is proportional to e^2, it is orthogonal to the four-velocity, and it depends on the rate of change

of the external force. This is the result that was first derived by Dirac [1]. (Dirac's original expression actually involved the rate of change of the acceleration vector on the right-hand side. The resulting equation gives rise to the well-known problem of runaway solutions. To avoid such unphysical behaviour we have submitted Dirac's equation to a reduction-of-order procedure whereby $da^\nu/d\tau$ is replaced with $m^{-1} df^\nu_{\rm ext}/d\tau$. This procedure is explained and justified, for example, in Refs. [11, 12], and further discussed in Sec. 24 below.)

To establish that the singular field exerts no force on the particle requires a careful analysis that is presented in the bulk of the paper. What really happens is that, because the particle is a monopole source for the electromagnetic field, the singular field is locally isotropic around the particle; it therefore exerts no force, but contributes to the particle's inertia and renormalizes its mass. In fact, one could do without a decomposition of the field into singular and regular solutions, and instead construct the force by using the retarded field and averaging it over a small sphere around the particle, as was done by Quinn and Wald [7]. In the body of this review we will use both methods and emphasize the equivalence of the results. We will, however, give some emphasis to the decomposition because it provides a compelling physical interpretation of the self-force as an interaction with a free electromagnetic field.

1.3 Green's functions in flat spacetime

To see how Eq. (1.5) can eventually be generalized to curved spacetimes, we introduce a new layer of mathematical formalism and show that the decomposition of the retarded potential into singular and regular pieces can be performed at the level of the Green's functions associated with Eq. (1.1). The retarded solution to the wave equation can be expressed as

$$A^\alpha_{\rm ret}(x) = \int G_{+\ \beta'}^{\ \alpha}(x,x') j^{\beta'}(x')\, dV', \tag{1.6}$$

in terms of the retarded Green's function $G_{+\ \beta'}^{\ \alpha}(x,x') = \delta^\alpha_{\beta'}\delta(t - t' - |\boldsymbol{x} - \boldsymbol{x'}|)/|\boldsymbol{x} - \boldsymbol{x'}|$. Here $x = (t, \boldsymbol{x})$ is an arbitrary field point, $x' = (t', \boldsymbol{x'})$ is a source point, and $dV' := d^4x'$; tensors at x are identified with unprimed indices, while primed indices refer to tensors at x'. Similarly, the advanced solution can be expressed as

$$A^\alpha_{\rm adv}(x) = \int G_{-\ \beta'}^{\ \alpha}(x,x') j^{\beta'}(x')\, dV', \tag{1.7}$$

in terms of the advanced Green's function $G_{-\ \beta'}^{\ \alpha}(x,x') = \delta^\alpha_{\beta'}\delta(t - t' + |\boldsymbol{x} - \boldsymbol{x'}|)/|\boldsymbol{x} - \boldsymbol{x'}|$. The retarded Green's function is zero whenever x lies outside of the future light cone of x', and $G_{+\ \beta'}^{\ \alpha}(x,x')$ is infinite at these points. On the other hand, the advanced Green's function is zero whenever x lies outside of the past light cone of x', and $G_{-\ \beta'}^{\ \alpha}(x,x')$ is infinite at these points. The retarded and advanced Green's functions satisfy the reciprocity relation

$$G^-_{\beta'\alpha}(x',x) = G^+_{\alpha\beta'}(x,x'); \tag{1.8}$$

this states that the retarded Green's function becomes the advanced Green's function (and vice versa) when x and x' are interchanged.

From the retarded and advanced Green's functions we can define a singular Green's function by

$$G_{{\rm S}\ \beta'}^{\ \alpha}(x,x') = \frac{1}{2}\Bigl[G_{+\ \beta'}^{\ \alpha}(x,x') + G_{-\ \beta'}^{\ \alpha}(x,x')\Bigr] \tag{1.9}$$

and a regular two-point function by

$$G_{{\rm R}\ \beta'}^{\ \alpha}(x,x') = G_{+\ \beta'}^{\ \alpha}(x,x') - G_{{\rm S}\ \beta'}^{\ \alpha}(x,x') = \frac{1}{2}\Bigl[G_{+\ \beta'}^{\ \alpha}(x,x') - G_{-\ \beta'}^{\ \alpha}(x,x')\Bigr]. \tag{1.10}$$

By virtue of Eq. (1.8) the singular Green's function is symmetric in its indices and arguments: $G^{\rm S}_{\beta'\alpha}(x',x) = G^{\rm S}_{\alpha\beta'}(x,x')$. The regular two-point function, on the other hand, is antisymmetric. The potential

$$A^\alpha_{\rm S}(x) = \int G_{{\rm S}\ \beta'}^{\ \alpha}(x,x') j^{\beta'}(x')\, dV' \tag{1.11}$$

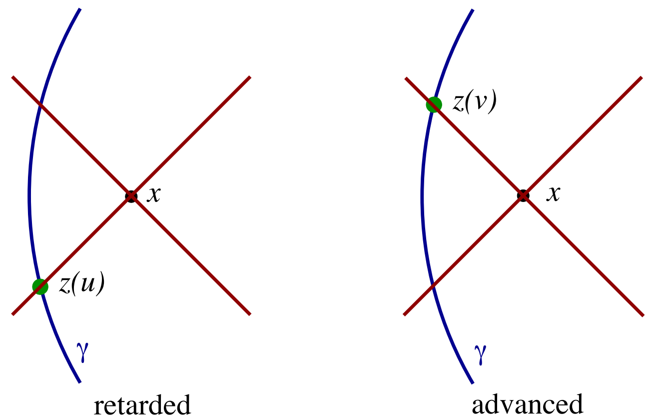

Figure 1: In flat spacetime, the retarded potential at x depends on the particle's state of motion at the retarded point $z(u)$ on the world line; the advanced potential depends on the state of motion at the advanced point $z(v)$.

satisfies the wave equation of Eq. (1.1) and is singular on the world line, while

$$A^\alpha_{\mathrm{R}}(x) = \int G_{\mathrm{R}}{}^{\alpha}{}_{\beta'}(x,x') j^{\beta'}(x')\, dV' \tag{1.12}$$

satisfies the homogeneous equation $\Box A^\alpha = 0$ and is well behaved on the world line.

Equation (1.6) implies that the retarded potential at x is generated by a single event in spacetime: the intersection of the world line and x's past light cone (see Fig. 1). We shall call this the *retarded point* associated with x and denote it $z(u)$; u is the *retarded time*, the value of the proper-time parameter at the retarded point. Similarly we find that the advanced potential of Eq. (1.7) is generated by the intersection of the world line and the future light cone of the field point x. We shall call this the *advanced point* associated with x and denote it $z(v)$; v is the *advanced time*, the value of the proper-time parameter at the advanced point.

1.4 Green's functions in curved spacetime

In a curved spacetime with metric $g_{\alpha\beta}$ the wave equation for the vector potential becomes

$$\Box A^\alpha - R^\alpha{}_\beta A^\beta = -4\pi j^\alpha, \tag{1.13}$$

where $\Box = g^{\alpha\beta}\nabla_\alpha\nabla_\beta$ is the covariant wave operator and $R_{\alpha\beta}$ is the spacetime's Ricci tensor; the Lorenz gauge conditions becomes $\nabla_\alpha A^\alpha = 0$, and ∇_α denotes covariant differentiation. Retarded and advanced Green's functions can be defined for this equation, and solutions to Eq. (1.13) take the same form as in Eqs. (1.6) and (1.7), except that dV' now stands for $\sqrt{-g(x')}\, d^4x'$.

The causal structure of the Green's functions is richer in curved spacetime: While in flat spacetime the retarded Green's function has support only on the future light cone of x', in curved spacetime its support extends *inside* the light cone as well; $G_{+}{}^{\alpha}{}_{\beta'}(x,x')$ is therefore nonzero when $x \in I^+(x')$, which denotes the chronological future of x'. This property reflects the fact that in curved spacetime, electromagnetic waves propagate not just at the speed of light, but at *all speeds smaller than or equal to the speed of light*; the delay is caused by an interaction between the radiation and the spacetime curvature. A direct implication of this property is that the retarded potential at x is now generated by the point charge during its entire history prior to the retarded time u associated with x: the potential depends on the particle's state of motion for all times $\tau \leq u$ (see Fig. 2).

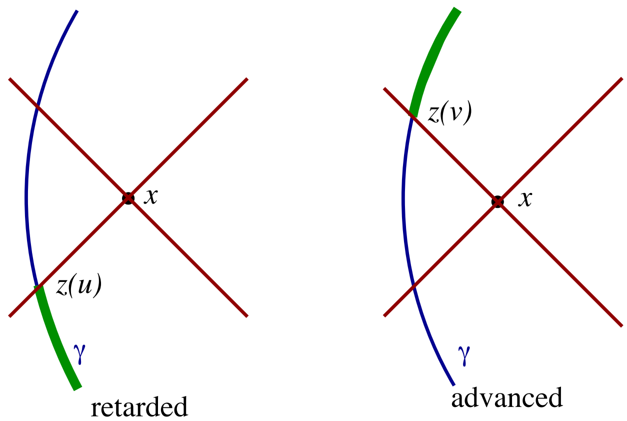

Figure 2: In curved spacetime, the retarded potential at x depends on the particle's history before the retarded time u; the advanced potential depends on the particle's history after the advanced time v.

Similar statements can be made about the advanced Green's function and the advanced solution to the wave equation. While in flat spacetime the advanced Green's function has support only on the past light cone of x', in curved spacetime its support extends inside the light cone, and $G_{-\ \beta'}^{\ \alpha}(x,x')$ is nonzero when $x \in I^{-}(x')$, which denotes the chronological past of x'. This implies that the advanced potential at x is generated by the point charge during its entire *future* history following the advanced time v associated with x: the potential depends on the particle's state of motion for all times $\tau \geq v$.

The physically relevant solution to Eq. (1.13) is obviously the retarded potential $A^{\alpha}_{\mathrm{ret}}(x)$, and as in flat spacetime, this diverges on the world line. The cause of this singular behaviour is still the pointlike nature of the source, and the presence of spacetime curvature does not change the fact that the potential diverges at the position of the particle. Once more this behaviour makes it difficult to figure out how the retarded field is supposed to act on the particle and determine its motion. As in flat spacetime we shall attempt to decompose the retarded solution into a singular part that exerts no force, and a regular part that produces the entire self-force.

To decompose the retarded Green's function into singular and regular parts is not a straightforward task in curved spacetime. The flat-spacetime definition for the singular Green's function, Eq. (1.9), cannot be adopted without modification: While the combination half-retarded plus half-advanced Green's functions does have the property of being symmetric, and while the resulting vector potential would be a solution to Eq. (1.13), this candidate for the singular Green's function would produce a self-force with an unacceptable dependence on the particle's future history. For suppose that we made this choice. Then the regular two-point function would be given by the combination half-retarded minus half-advanced Green's functions, just as in flat spacetime. The resulting potential would satisfy the homogeneous wave equation, and it would be regular on the world line, but it would also depend on the particle's entire history, both past (through the retarded Green's function) and future (through the advanced Green's function). More precisely stated, we would find that the regular potential at x depends on the particle's state of motion at all times τ outside the interval $u < \tau < v$; in the limit where x approaches the world line, this interval shrinks to nothing, and we would find that the regular potential is generated by the complete history of the particle. A self-force constructed from this potential would be highly noncausal, and we are compelled to reject these definitions for the singular and regular Green's functions.

The proper definitions were identified by Detweiler and Whiting [13], who proposed the following generalization to Eq. (1.9):

$$G_{\mathrm{S}\ \beta'}^{\ \alpha}(x,x') = \frac{1}{2}\Big[G_{+\ \beta'}^{\ \alpha}(x,x') + G_{-\ \beta'}^{\ \alpha}(x,x') - H^{\alpha}_{\ \beta'}(x,x')\Big]. \tag{1.14}$$

The two-point function $H^{\alpha}_{\ \beta'}(x,x')$ is introduced specifically to cure the pathology described in the preceding

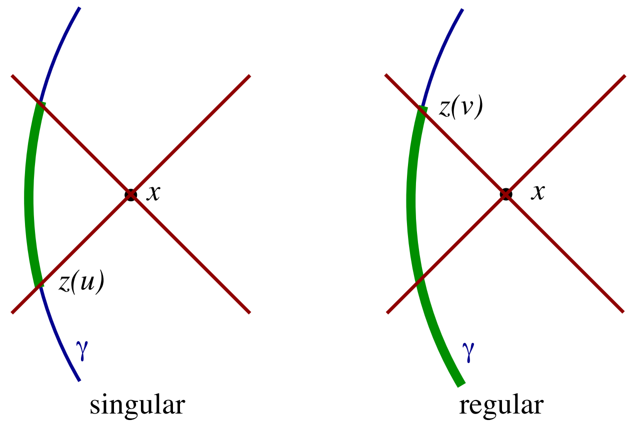

Figure 3: In curved spacetime, the singular potential at x depends on the particle's history during the interval $u \le \tau \le v$; for the regular potential the relevant interval is $-\infty < \tau \le v$.

paragraph. It is symmetric in its indices and arguments, so that $G^{\rm S}_{\alpha\beta'}(x,x')$ will be also (since the retarded and advanced Green's functions are still linked by a reciprocity relation); and it is a solution to the homogeneous wave equation, $\Box H^{\alpha}_{\ \beta'}(x,x') - R^{\alpha}_{\ \gamma}(x) H^{\gamma}_{\ \beta'}(x,x') = 0$, so that the singular, retarded, and advanced Green's functions will all satisfy the same wave equation. Furthermore, and this is its key property, the two-point function is defined to agree with the advanced Green's function when x is in the chronological past of x': $H^{\alpha}_{\ \beta'}(x,x') = G_{-\ \beta'}^{\ \alpha}(x,x')$ when $x \in I^-(x')$. This ensures that $G_{\rm S\ \beta'}^{\ \alpha}(x,x')$ vanishes when x is in the chronological past of x'. In fact, reciprocity implies that $H^{\alpha}_{\ \beta'}(x,x')$ will also agree with the *retarded* Green's function when x is in the chronological future of x', and it follows that the symmetric Green's function vanishes also when x is in the chronological future of x'.

The potential $A^{\alpha}_{\rm S}(x)$ constructed from the singular Green's function can now be seen to depend on the particle's state of motion at times τ restricted to the interval $u \le \tau \le v$ (see Fig. 3). Because this potential satisfies Eq. (1.13), it is just as singular as the retarded potential in the vicinity of the world line. And because the singular Green's function is symmetric in its arguments, the singular potential can be shown to exert no force on the charged particle. (This requires a lengthy analysis that will be presented in the bulk of the paper.)

The Detweiler-Whiting [13] definition for the regular two-point function is then

$$G_{\rm R\ \beta'}^{\ \alpha}(x,x') = G_{+\ \beta'}^{\ \alpha}(x,x') - G_{\rm S\ \beta'}^{\ \alpha}(x,x') = \frac{1}{2}\Big[G_{+\ \beta'}^{\ \alpha}(x,x') - G_{-\ \beta'}^{\ \alpha}(x,x') + H^{\alpha}_{\ \beta'}(x,x')\Big]. \qquad (1.15)$$

The potential $A^{\alpha}_{\rm R}(x)$ constructed from this depends on the particle's state of motion at all times τ prior to the advanced time v: $\tau \le v$. Because this potential satisfies the homogeneous wave equation, it is well behaved on the world line and its action on the point charge is well defined. And because the singular potential $A^{\alpha}_{\rm S}(x)$ can be shown to exert no force on the particle, we conclude that $A^{\alpha}_{\rm R}(x)$ alone is responsible for the self-force.

From the regular potential we form an electromagnetic field tensor $F^{\rm R}_{\alpha\beta} = \nabla_\alpha A^{\rm R}_\beta - \nabla_\beta A^{\rm R}_\alpha$ and the curved-spacetime generalization to Eq. (1.4) is

$$m a_\mu = f^{\rm ext}_\mu + e F^{\rm R}_{\mu\nu} u^\nu, \qquad (1.16)$$

where $u^\mu = dz^\mu/d\tau$ is again the charge's four-velocity, but $a^\mu = Du^\mu/d\tau$ is now its covariant acceleration.

1.5 World line and retarded coordinates

To flesh out the ideas contained in the preceding subsection we add yet another layer of mathematical formalism and construct a convenient coordinate system to chart a neighbourhood of the particle's world

line. In the next subsection we will display explicit expressions for the retarded, singular, and regular fields of a point electric charge.

Let γ be the world line of a point particle in a curved spacetime. It is described by parametric relations $z^\mu(\tau)$ in which τ is proper time. Its tangent vector is $u^\mu = dz^\mu/d\tau$ and its acceleration is $a^\mu = Du^\mu/d\tau$; we shall also encounter $\dot{a}^\mu := Da^\mu/d\tau$.

On γ we erect an orthonormal basis that consists of the four-velocity u^μ and three spatial vectors e^μ_a labelled by a frame index $a = (1,2,3)$. These vectors satisfy the relations $g_{\mu\nu}u^\mu u^\nu = -1$, $g_{\mu\nu}u^\mu e^\nu_a = 0$, and $g_{\mu\nu}e^\mu_a e^\nu_b = \delta_{ab}$. We take the spatial vectors to be Fermi-Walker transported on the world line: $De^\mu_a/d\tau = a_a u^\mu$, where

$$a_a(\tau) = a_\mu e^\mu_a \tag{1.17}$$

are frame components of the acceleration vector; it is easy to show that Fermi-Walker transport preserves the orthonormality of the basis vectors. We shall use the tetrad to decompose various tensors evaluated on the world line. An example was already given in Eq. (1.17) but we shall also encounter frame components of the Riemann tensor,

$$R_{a0b0}(\tau) = R_{\mu\lambda\nu\rho}e^\mu_a u^\lambda e^\nu_b u^\rho, \qquad R_{a0bc}(\tau) = R_{\mu\lambda\nu\rho}e^\mu_a u^\lambda e^\nu_b e^\rho_c, \qquad R_{abcd}(\tau) = R_{\mu\lambda\nu\rho}e^\mu_a e^\lambda_b e^\nu_c e^\rho_d, \tag{1.18}$$

as well as frame components of the Ricci tensor,

$$R_{00}(\tau) = R_{\mu\nu}u^\mu u^\nu, \qquad R_{a0}(\tau) = R_{\mu\nu}e^\mu_a u^\nu, \qquad R_{ab}(\tau) = R_{\mu\nu}e^\mu_a e^\nu_b. \tag{1.19}$$

We shall use $\delta_{ab} = \mathrm{diag}(1,1,1)$ and its inverse $\delta^{ab} = \mathrm{diag}(1,1,1)$ to lower and raise frame indices, respectively.

Consider a point x in a neighbourhood of the world line γ. We assume that x is sufficiently close to the world line that a unique geodesic links x to any neighbouring point z on γ. The two-point function $\sigma(x,z)$, known as *Synge's world function* [14], is numerically equal to half the squared geodesic distance between z and x; it is positive if x and z are spacelike related, negative if they are timelike related, and $\sigma(x,z)$ is zero if x and z are linked by a null geodesic. We denote its gradient $\partial\sigma/\partial z^\mu$ by $\sigma_\mu(x,z)$, and $-\sigma^\mu$ gives a meaningful notion of a separation vector (pointing from z to x).

To construct a coordinate system in this neighbourhood we locate the unique point $x' := z(u)$ on γ which is linked to x by a future-directed null geodesic (this geodesic is directed from x' to x); we shall refer to x' as the *retarded point* associated with x, and u will be called the *retarded time*. To tensors at x' we assign indices α', β', …; this will distinguish them from tensors at a generic point $z(\tau)$ on the world line, to which we have assigned indices μ, ν, …. We have $\sigma(x,x') = 0$ and $-\sigma^{\alpha'}(x,x')$ is a null vector that can be interpreted as the separation between x' and x.

The *retarded coordinates* of the point x are $(u, \hat{x}^a)$, where $\hat{x}^a = -e^a_{\alpha'}\sigma^{\alpha'}$ are the frame components of the separation vector. They come with a straightforward interpretation (see Fig. 4). The invariant quantity

$$r := \sqrt{\delta_{ab}\hat{x}^a\hat{x}^b} = u_{\alpha'}\sigma^{\alpha'} \tag{1.20}$$

is an affine parameter on the null geodesic that links x to x'; it can be loosely interpreted as the time delay between x and x' as measured by an observer moving with the particle. This therefore gives a meaningful notion of distance between x and the retarded point, and we shall call r the *retarded distance* between x and the world line. The unit vector

$$\Omega^a = \hat{x}^a/r \tag{1.21}$$

is constant on the null geodesic that links x to x'. Because Ω^a is a different constant on each null geodesic that emanates from x', keeping u fixed and varying Ω^a produces a congruence of null geodesics that generate the future light cone of the point x' (the congruence is hypersurface orthogonal). Each light cone can thus be labelled by its retarded time u, each generator on a given light cone can be labelled by its direction vector Ω^a, and each point on a given generator can be labelled by its retarded distance r. We therefore have a good coordinate system in a neighbourhood of γ.

To tensors at x we assign indices α, β, …. These tensors will be decomposed in a tetrad (e^α_0, e^α_a) that is constructed as follows: Given x we locate its associated retarded point x' on the world line, as well as the null geodesic that links these two points; we then take the tetrad $(u^{\alpha'}, e^{\alpha'}_a)$ at x' and parallel transport it to x along the null geodesic to obtain (e^α_0, e^α_a).

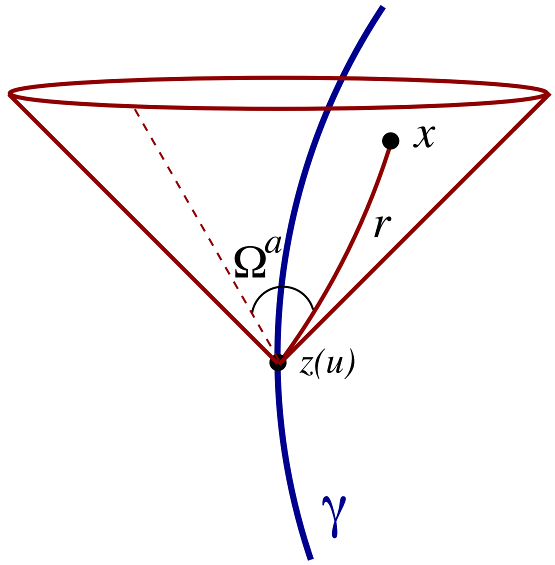

Figure 4: Retarded coordinates of a point x relative to a world line γ. The retarded time u selects a particular null cone, the unit vector $\Omega^a := \hat{x}^a/r$ selects a particular generator of this null cone, and the retarded distance r selects a particular point on this generator.

1.6 Retarded, singular, and regular electromagnetic fields of a point electric charge

The retarded solution to Eq. (1.13) is

$$A^\alpha(x) = e\int_\gamma G_{+\mu}^{\ \alpha}(x,z)u^\mu\,d\tau, \tag{1.22}$$

where the integration is over the world line of the point electric charge. Because the retarded solution is the physically relevant solution to the wave equation, it will not be necessary to put a label 'ret' on the vector potential.

From the vector potential we form the electromagnetic field tensor $F_{\alpha\beta}$, which we decompose in the tetrad (e^α_0, e^α_a) introduced at the end of Sec. 1.5. We then express the frame components of the field tensor in retarded coordinates, in the form of an expansion in powers of r. This gives

$$\begin{aligned}
F_{a0}(u,r,\Omega^a) &:= F_{\alpha\beta}(x)e^\alpha_a(x)e^\beta_0(x) \\
&= \frac{e}{r^2}\Omega_a - \frac{e}{r}\bigl(a_a - a_b\Omega^b\Omega_a\bigr) + \frac{1}{3}eR_{b0c0}\Omega^b\Omega^c\Omega_a - \frac{1}{6}e\bigl(5R_{a0b0}\Omega^b + R_{ab0c}\Omega^b\Omega^c\bigr) \\
&\quad + \frac{1}{12}e\bigl(5R_{00} + R_{bc}\Omega^b\Omega^c + R\bigr)\Omega_a + \frac{1}{3}eR_{a0} - \frac{1}{6}eR_{ab}\Omega^b + F^{\rm tail}_{a0} + O(r), \qquad (1.23) \\
F_{ab}(u,r,\Omega^a) &:= F_{\alpha\beta}(x)e^\alpha_a(x)e^\beta_b(x) \\
&= \frac{e}{r}\bigl(a_a\Omega_b - \Omega_a a_b\bigr) + \frac{1}{2}e\bigl(R_{a0bc} - R_{b0ac} + R_{a0c0}\Omega_b - \Omega_a R_{b0c0}\bigr)\Omega^c \\
&\quad - \frac{1}{2}e\bigl(R_{a0}\Omega_b - \Omega_a R_{b0}\bigr) + F^{\rm tail}_{ab} + O(r), \qquad (1.24)
\end{aligned}$$

where

$$F^{\rm tail}_{a0} = F^{\rm tail}_{\alpha'\beta'}(x')e^{\alpha'}_a u^{\beta'}, \qquad F^{\rm tail}_{ab} = F^{\rm tail}_{\alpha'\beta'}(x')e^{\alpha'}_a e^{\beta'}_b \tag{1.25}$$

are the frame components of the "tail part" of the field, which is given by

$$F^{\rm tail}_{\alpha'\beta'}(x') = 2e\int_{-\infty}^{u^-}\nabla_{[\alpha'}G_{+\beta']\mu}(x',z)u^\mu\,d\tau. \tag{1.26}$$

In these expressions, all tensors (or their frame components) are evaluated at the retarded point $x' := z(u)$ associated with x; for example, $a_a := a_a(u) := a_{\alpha'} e^{\alpha'}_a$. The tail part of the electromagnetic field tensor is written as an integral over the portion of the world line that corresponds to the interval $-\infty < \tau \leq u^- := u - 0^+$; this represents the past history of the particle. The integral is cut short at u^- to avoid the singular behaviour of the retarded Green's function when $z(\tau)$ coincides with x'; the portion of the Green's function involved in the tail integral is smooth, and the singularity at coincidence is completely accounted for by the other terms in Eqs. (1.23) and (1.24).

The expansion of $F_{\alpha\beta}(x)$ near the world line does indeed reveal many singular terms. We first recognize terms that diverge when $r \to 0$; for example the Coulomb field F_{a0} diverges as r^{-2} when we approach the world line. But there are also terms that, though they stay bounded in the limit, possess a directional ambiguity at $r = 0$; for example F_{ab} contains a term proportional to $R_{a0bc}\Omega^c$ whose limit depends on the direction of approach.

This singularity structure is perfectly reproduced by the singular field $F^{\rm S}_{\alpha\beta}$ obtained from the potential

$$A^\alpha_{\rm S}(x) = e \int_\gamma G^{\ \alpha}_{{\rm S}\ \mu}(x, z) u^\mu \, d\tau, \tag{1.27}$$

where $G^{\ \alpha}_{{\rm S}\ \mu}(x, z)$ is the singular Green's function of Eq. (1.14). Near the world line the singular field is given by

$$\begin{aligned}
F^{\rm S}_{a0}(u, r, \Omega^a) &:= F^{\rm S}_{\alpha\beta}(x) e^\alpha_a(x) e^\beta_0(x) \\
&= \frac{e}{r^2}\Omega_a - \frac{e}{r}\bigl(a_a - a_b\Omega^b\Omega_a\bigr) - \frac{2}{3}e\dot{a}_a + \frac{1}{3}eR_{b0c0}\Omega^b\Omega^c\Omega_a - \frac{1}{6}e\bigl(5R_{a0b0}\Omega^b + R_{ab0c}\Omega^b\Omega^c\bigr) \\
&\quad + \frac{1}{12}e\bigl(5R_{00} + R_{bc}\Omega^b\Omega^c + R\bigr)\Omega_a - \frac{1}{6}eR_{ab}\Omega^b + O(r),
\end{aligned} \tag{1.28}$$

$$\begin{aligned}
F^{\rm S}_{ab}(u, r, \Omega^a) &:= F^{\rm S}_{\alpha\beta}(x) e^\alpha_a(x) e^\beta_b(x) \\
&= \frac{e}{r}\bigl(a_a\Omega_b - \Omega_a a_b\bigr) + \frac{1}{2}e\bigl(R_{a0bc} - R_{b0ac} + R_{a0c0}\Omega_b - \Omega_a R_{b0c0}\bigr)\Omega^c \\
&\quad - \frac{1}{2}e\bigl(R_{a0}\Omega_b - \Omega_a R_{b0}\bigr) + O(r).
\end{aligned} \tag{1.29}$$

Comparison of these expressions with Eqs. (1.23) and (1.24) does indeed reveal that all singular terms are shared by both fields.

The difference between the retarded and singular fields defines the regular field $F^{\rm R}_{\alpha\beta}(x)$. Its frame components are

$$F^{\rm R}_{a0} = \frac{2}{3}e\dot{a}_a + \frac{1}{3}eR_{a0} + F^{\rm tail}_{a0} + O(r), \tag{1.30}$$

$$F^{\rm R}_{ab} = F^{\rm tail}_{ab} + O(r), \tag{1.31}$$

and at x' the regular field becomes

$$F^{\rm R}_{\alpha'\beta'} = 2eu_{[\alpha'}\bigl(g_{\beta']\gamma'} + u_{\beta']}u_{\gamma'}\bigr)\biggl(\frac{2}{3}\dot{a}^{\gamma'} + \frac{1}{3}R^{\gamma'}_{\ \ \delta'}u^{\delta'}\biggr) + F^{\rm tail}_{\alpha'\beta'}, \tag{1.32}$$

where $\dot{a}^{\gamma'} = Da^{\gamma'}/d\tau$ is the rate of change of the acceleration vector, and where the tail term was given by Eq. (1.26). We see that $F^{\rm R}_{\alpha\beta}(x)$ is a regular tensor field, even on the world line.

1.7 Motion of an electric charge in curved spacetime

We have argued in Sec. 1.4 that the self-force acting on a point electric charge is produced by the regular field, and that the charge's equations of motion should take the form of $ma_\mu = f^{\rm ext}_\mu + eF^{\rm R}_{\mu\nu}u^\nu$, where $f^{\rm ext}_\mu$ is an external force also acting on the particle. Substituting Eq. (1.32) gives

$$ma^\mu = f^\mu_{\rm ext} + e^2\bigl(\delta^\mu_{\ \nu} + u^\mu u_\nu\bigr)\biggl(\frac{2}{3m}\frac{Df^\nu_{\rm ext}}{d\tau} + \frac{1}{3}R^\nu_{\ \lambda}u^\lambda\biggr) + 2e^2 u_\nu \int_{-\infty}^{\tau^-} \nabla^{[\mu}G^{\ \nu]}_{+\ \lambda'}\bigl(z(\tau), z(\tau')\bigr)u^{\lambda'}\, d\tau', \tag{1.33}$$

in which all tensors are evaluated at $z(\tau)$, the current position of the particle on the world line. The primed indices in the tail integral refer to a point $z(\tau')$ which represents a prior position; the integration is cut short at $\tau' = \tau^- := \tau - 0^+$ to avoid the singular behaviour of the retarded Green's function at coincidence. To get Eq. (1.33) we have reduced the order of the differential equation by replacing $\dot{a}^\nu$ with $m^{-1}\dot{f}^\nu_{\rm ext}$ on the right-hand side; this procedure was explained at the end of Sec. 1.2.

Equation (1.33) is the result that was first derived by DeWitt and Brehme [4] and later corrected by Hobbs [5]. (The original version of the equation did not include the Ricci-tensor term.) In flat spacetime the Ricci tensor is zero, the tail integral disappears (because the Green's function vanishes everywhere within the domain of integration), and Eq. (1.33) reduces to Dirac's result of Eq. (1.5). In curved spacetime the self-force does not vanish even when the electric charge is moving freely, in the absence of an external force: it is then given by the tail integral, which represents radiation emitted earlier and coming back to the particle after interacting with the spacetime curvature. This delayed action implies that in general, the self-force is nonlocal in time: it depends not only on the current state of motion of the particle, but also on its past history. Lest this behaviour should seem mysterious, it may help to keep in mind that the physical process that leads to Eq. (1.33) is simply an interaction between the charge and a free electromagnetic field $F^{\rm R}_{\alpha\beta}$; it is this field that carries the information about the charge's past.

1.8 Motion of a scalar charge in curved spacetime

The dynamics of a point scalar charge can be formulated in a way that stays fairly close to the electromagnetic theory. The particle's charge q produces a scalar field $\Phi(x)$ which satisfies a wave equation

$$\bigl(\Box - \xi R\bigr)\Phi = -4\pi\mu \tag{1.34}$$

that is very similar to Eq. (1.13). Here, R is the spacetime's Ricci scalar, and ξ is an arbitrary coupling constant; the scalar charge density $\mu(x)$ is given by a four-dimensional Dirac functional supported on the particle's world line γ. The retarded solution to the wave equation is

$$\Phi(x) = q\int_\gamma G_+(x,z)\, d\tau, \tag{1.35}$$

where $G_+(x,z)$ is the retarded Green's function associated with Eq. (1.34). The field exerts a force on the particle, whose equations of motion are

$$ma^\mu = q\bigl(g^{\mu\nu} + u^\mu u^\nu\bigr)\nabla_\nu \Phi, \tag{1.36}$$

where m is the particle's mass; this equation is very similar to the Lorentz-force law. But the dynamics of a scalar charge comes with a twist: If Eqs. (1.34) and (1.36) are to follow from a variational principle, *the particle's mass should not be expected to be a constant of the motion.* It is found instead to satisfy the differential equation

$$\frac{dm}{d\tau} = -q u^\mu \nabla_\mu \Phi, \tag{1.37}$$

and in general m will vary with proper time. This phenomenon is linked to the fact that a scalar field has zero spin: the particle can radiate monopole waves and the radiated energy can come at the expense of the rest mass.

The scalar field of Eq. (1.35) diverges on the world line and its singular part $\Phi_{\rm S}(x)$ must be removed before Eqs. (1.36) and (1.37) can be evaluated. This procedure produces the regular field $\Phi_{\rm R}(x)$, and it is this field (which satisfies the homogeneous wave equation) that gives rise to a self-force. The gradient of the regular field takes the form of

$$\nabla_\mu \Phi_{\rm R} = -\frac{1}{12}(1-6\xi)qRu_\mu + q\bigl(g_{\mu\nu} + u_\mu u_\nu\bigr)\biggl(\frac{1}{3}\dot{a}^\nu + \frac{1}{6}R^\nu_{\ \lambda}u^\lambda\biggr) + \Phi^{\rm tail}_\mu \tag{1.38}$$

when it is evaluated on the world line. The last term is the tail integral

$$\Phi^{\rm tail}_\mu = q\int_{-\infty}^{\tau^-} \nabla_\mu G_+\bigl(z(\tau),z(\tau')\bigr)\, d\tau', \tag{1.39}$$

and this brings the dependence on the particle's past.

Substitution of Eq. (1.38) into Eqs. (1.36) and (1.37) gives the equations of motion of a point scalar charge. (At this stage we introduce an external force $f^\mu_{\rm ext}$ and reduce the order of the differential equation.) The acceleration is given by

$$ma^\mu = f^\mu_{\rm ext} + q^2\bigl(\delta^\mu_{\ \nu} + u^\mu u_\nu\bigr)\biggl[\frac{1}{3m}\frac{Df^\nu_{\rm ext}}{d\tau} + \frac{1}{6}R^\nu_{\ \lambda}u^\lambda + \int_{-\infty}^{\tau^-}\nabla^\nu G_+\bigl(z(\tau),z(\tau')\bigr)\,d\tau'\biggr] \tag{1.40}$$

and the mass changes according to

$$\frac{dm}{d\tau} = -\frac{1}{12}(1-6\xi)q^2R - q^2u^\mu\int_{-\infty}^{\tau^-}\nabla_\mu G_+\bigl(z(\tau),z(\tau')\bigr)\,d\tau'. \tag{1.41}$$

These equations were first derived by Quinn [8]. (His analysis was restricted to a minimally coupled scalar field, so that $\xi = 0$ in his expressions. We extended Quinn's results to an arbitrary coupling counstant for this review.)

In flat spacetime the Ricci-tensor term and the tail integral disappear and Eq. (1.40) takes the form of Eq. (1.5) with $q^2/(3m)$ replacing the factor of $2e^2/(3m)$. In this simple case Eq. (1.41) reduces to $dm/d\tau = 0$ and the mass is in fact a constant. This property remains true in a conformally flat spacetime when the wave equation is conformally invariant ($\xi = 1/6$): in this case the Green's function possesses only a light-cone part and the right-hand side of Eq. (1.41) vanishes. In generic situations the mass of a point scalar charge will vary with proper time.

1.9 Motion of a point mass, or a small body, in a background spacetime

The case of a point mass moving in a specified background spacetime presents itself with a serious conceptual challenge, as the fundamental equations of the theory are nonlinear and the very notion of a "point mass" is somewhat misguided. Nevertheless, to the extent that the perturbation $h_{\alpha\beta}(x)$ created by the point mass can be considered to be "small", the problem can be formulated in close analogy with what was presented before.

We take the metric $g_{\alpha\beta}$ of the background spacetime to be a solution of the Einstein field equations in vacuum. (We impose this condition globally.) We describe the gravitational perturbation produced by a point particle of mass m in terms of trace-reversed potentials $\gamma_{\alpha\beta}$ defined by

$$\gamma_{\alpha\beta} = h_{\alpha\beta} - \frac{1}{2}\bigl(g^{\gamma\delta}h_{\gamma\delta}\bigr)g_{\alpha\beta}, \tag{1.42}$$

where $h_{\alpha\beta}$ is the difference between $\mathsf{g}_{\alpha\beta}$, the actual metric of the perturbed spacetime, and $g_{\alpha\beta}$. The potentials satisfy the wave equation

$$\Box\gamma^{\alpha\beta} + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}\gamma^{\gamma\delta} = -16\pi T^{\alpha\beta} + O(m^2) \tag{1.43}$$

together with the Lorenz gauge condition $\gamma^{\alpha\beta}_{\ \ \ ;\beta} = 0$. Here and below, covariant differentiation refers to a connection that is compatible with the background metric, $\Box = g^{\alpha\beta}\nabla_\alpha\nabla_\beta$ is the wave operator for the background spacetime, and $T^{\alpha\beta}$ is the energy-momentum tensor of the point mass; this is given by a Dirac distribution supported on the particle's world line γ. The retarded solution is

$$\gamma^{\alpha\beta}(x) = 4m\int_\gamma G^{\ \alpha\beta}_{+\ \mu\nu}(x,z)u^\mu u^\nu\,d\tau + O(m^2), \tag{1.44}$$

where $G^{\ \alpha\beta}_{+\ \mu\nu}(x,z)$ is the retarded Green's function associated with Eq. (1.43). The perturbation $h_{\alpha\beta}(x)$ can be recovered by inverting Eq. (1.42).

Equations of motion for the point mass can be obtained by formally demanding that the motion be geodesic in the perturbed spacetime with metric $\mathsf{g}_{\alpha\beta} = g_{\alpha\beta} + h_{\alpha\beta}$. After a mapping to the background spacetime, the equations of motion take the form of

$$a^\mu = -\frac{1}{2}\bigl(g^{\mu\nu} + u^\mu u^\nu\bigr)\bigl(2h_{\nu\lambda;\rho} - h_{\lambda\rho;\nu}\bigr)u^\lambda u^\rho + O(m^2). \tag{1.45}$$

The acceleration is thus proportional to m; in the test-mass limit the world line of the particle is a geodesic of the background spacetime.

We now remove $h^{\rm S}_{\alpha\beta}(x)$ from the retarded perturbation and postulate that it is the regular field $h^{\rm R}_{\alpha\beta}(x)$ that should act on the particle. (Note that $\gamma^{\rm S}_{\alpha\beta}$ satisfies the same wave equation as the retarded potentials, but that $\gamma^{\rm R}_{\alpha\beta}$ is a free gravitational field that satisfies the homogeneous wave equation.) On the world line we have

$$h^{\rm R}_{\mu\nu;\lambda} = -4m\Bigl(u_{(\mu}R_{\nu)\rho\lambda\xi} + R_{\mu\rho\nu\xi}u_\lambda\Bigr)u^\rho u^\xi + h^{\rm tail}_{\mu\nu\lambda}, \tag{1.46}$$

where the tail term is given by

$$h^{\rm tail}_{\mu\nu\lambda} = 4m\int_{-\infty}^{\tau^-} \nabla_\lambda\biggl(G_{+\mu\nu\mu'\nu'} - \frac{1}{2}g_{\mu\nu}G_{+\ \ \rho\mu'\nu'}^{\ \ \rho}\biggr)\bigl(z(\tau),z(\tau')\bigr)u^{\mu'}u^{\nu'}\,d\tau'. \tag{1.47}$$

When Eq. (1.46) is substituted into Eq. (1.45) we find that the terms that involve the Riemann tensor cancel out, and we are left with

$$a^\mu = -\frac{1}{2}\bigl(g^{\mu\nu}+u^\mu u^\nu\bigr)\bigl(2h^{\rm tail}_{\nu\lambda\rho} - h^{\rm tail}_{\lambda\rho\nu}\bigr)u^\lambda u^\rho + O(m^2). \tag{1.48}$$

Only the tail integral appears in the final form of the equations of motion. It involves the current position $z(\tau)$ of the particle, at which all tensors with unprimed indices are evaluated, as well as all prior positions $z(\tau')$, at which tensors with primed indices are evaluated. As before the integral is cut short at $\tau' = \tau^- := \tau - 0^+$ to avoid the singular behaviour of the retarded Green's function at coincidence.

The equations of motion of Eq. (1.48) were first derived by Mino, Sasaki, and Tanaka [6], and then reproduced with a different analysis by Quinn and Wald [7]. They are now known as the MiSaTaQuWa equations of motion. As noted by these authors, the MiSaTaQuWa equation has the appearance of the geodesic equation in a metric $g_{\alpha\beta} + h^{\rm tail}_{\alpha\beta}$. Detweiler and Whiting [13] have contributed the more compelling interpretation that the motion is actually geodesic in a spacetime with metric $g_{\alpha\beta} + h^{\rm R}_{\alpha\beta}$. The distinction is important: Unlike the first version of the metric, the Detweiler-Whiting metric is regular on the world line and satisfies the Einstein field equations in vacuum; and because it is a solution to the field equations, it can be viewed as a physical metric — specifically, the metric of the background spacetime perturbed by a free field produced by the particle at an earlier stage of its history.

While Eq. (1.48) does indeed give the correct equations of motion for a small mass m moving in a background spacetime with metric $g_{\alpha\beta}$, the derivation outlined here leaves much to be desired — to what extent should we trust an analysis based on the existence of a point mass? As a partial answer to this question, Mino, Sasaki, and Tanaka [6] produced an alternative derivation of their result, which involved a small nonrotating black hole instead of a point mass. In this alternative derivation, the metric of the black hole perturbed by the tidal gravitational field of the external universe is matched to the metric of the background spacetime perturbed by the moving black hole. Demanding that this metric be a solution to the vacuum field equations determines the motion of the black hole: it must move according to Eq. (1.48). This alternative derivation (which was given a different implementation in Ref. [15]) is entirely free of singularities (except deep within the black hole), and it suggests that the MiSaTaQuWa equations can be trusted to describe the motion of any gravitating body in a curved background spacetime (so long as the body's internal structure can be ignored). This derivation, however, was limited to the case of a non-rotating black hole, and it relied on a number of unjustified and sometimes unstated assumptions [16–18]. The conclusion was made firm by the more rigorous analysis of Gralla and Wald [16] (as extended by Pound [17]), who showed that the MiSaTaQuWa equations apply to any sufficiently compact body of arbitrary internal structure.

It is important to understand that unlike Eqs. (1.33) and (1.40), which are true tensorial equations, Eq. (1.48) reflects a specific choice of coordinate system and its form would not be preserved under a coordinate transformation. In other words, *the MiSaTaQuWa equations are not gauge invariant*, and they depend upon the Lorenz gauge condition $\gamma^{\alpha\beta}{}_{;\beta} = O(m^2)$. Barack and Ori [19] have shown that under a coordinate transformation of the form $x^\alpha \to x^\alpha + \xi^\alpha$, where x^α are the coordinates of the background spacetime and ξ^α is a smooth vector field of order m, the particle's acceleration changes according to $a^\mu \to a^\mu + a[\xi]^\mu$, where

$$a[\xi]^\mu = \bigl(\delta^\mu_{\ \nu} + u^\mu u_\nu\bigr)\biggl(\frac{D^2\xi^\nu}{d\tau^2} + R^\nu_{\ \rho\omega\lambda}u^\rho\xi^\omega u^\lambda\biggr) \tag{1.49}$$

is the "gauge acceleration"; $D^2\xi^\nu/d\tau^2 = (\xi^\nu{}_{;\mu}u^\mu)_{;\rho}u^\rho$ is the second covariant derivative of ξ^ν in the direction of the world line. This implies that the particle's acceleration can be altered at will by a gauge transformation; ξ^α could even be chosen so as to produce $a^\mu = 0$, making the motion geodesic after all. This observation provides a dramatic illustration of the following point: *The MiSaTaQuWa equations of motion are not gauge invariant and they cannot by themselves produce a meaningful answer to a well-posed physical question; to obtain such answers it is necessary to combine the equations of motion with the metric perturbation $h_{\alpha\beta}$ so as to form gauge-invariant quantities that will correspond to direct observables.* This point is very important and cannot be over-emphasized.

The gravitational self-force possesses a physical significance that is not shared by its scalar and electromagnetic analogues, because the motion of a small body in the strong gravitational field of a much larger body is a problem of direct relevance to gravitational-wave astronomy. Indeed, *extreme-mass-ratio inspirals*, involving solar-mass compact objects moving around massive black holes of the sort found in galactic cores, have been identified as promising sources of low-frequency gravitational waves for space-based interferometric detectors such as the proposed Laser Interferometer Space Antenna (LISA [20]). These systems involve highly eccentric, nonequatorial, and relativistic orbits around rapidly rotating black holes, and the waves produced by such orbital motions are rich in information concerning the strongest gravitational fields in the Universe. This information will be extractable from the LISA data stream, but the extraction depends on sophisticated data-analysis strategies that require a detailed and accurate modeling of the source. This modeling involves formulating the equations of motion for the small body in the field of the rotating black hole, as well as a consistent incorporation of the motion into a wave-generation formalism. In short, the extraction of this wealth of information relies on a successful evaluation of the gravitational self-force.

The finite-mass corrections to the orbital motion are important. For concreteness, let us assume that the orbiting body is a black hole of mass $m = 10\ M_\odot$ and that the central black hole has a mass $M = 10^6\ M_\odot$. Let us also assume that the small black hole is in the deep field of the large hole, near the innermost stable circular orbit, so that its orbital period P is of the order of minutes. The gravitational waves produced by the orbital motion have frequencies f of the order of the mHz, which is well within LISA's frequency band. The radiative losses drive the orbital motion toward a final plunge into the large black hole; this occurs over a radiation-reaction timescale $(M/m)P$ of the order of a year, during which the system will go through a number of wave cycles of the order of $M/m = 10^5$. *The role of the gravitational self-force is precisely to describe this orbital evolution toward the final plunge.* While at any given time the self-force provides fractional corrections of order $m/M = 10^{-5}$ to the motion of the small black hole, these build up over a number of orbital cycles of order $M/m = 10^5$ to produce a large cumulative effect. As will be discussed in some detail in Sec. 2.6, the gravitational self-force is important, because it drives large secular changes in the orbital motion of an extreme-mass-ratio binary.

1.10 Case study: static electric charge in Schwarzschild spacetime

One of the first self-force calculations ever performed for a curved spacetime was presented by Smith and Will [21]. They considered an electric charge e held in place at position $r = r_0$ outside a Schwarzschild black hole of mass M. Such a static particle must be maintained in position with an external force that compensates for the black hole's attraction. For a particle without electric charge this force is directed outward, and its radial component in Schwarzschild coordinates is given by $f^r_{\rm ext} = \frac{1}{2}mf'$, where m is the particle's mass, $f := 1 - 2M/r_0$ is the usual metric factor, and a prime indicates differentiation with respect to r_0, so that $f' = 2M/r_0^2$. Smith and Will found that for a particle of charge e, the external force is given instead by $f^r_{\rm ext} = \frac{1}{2}mf' - e^2Mf^{1/2}/r_0^3$. The second term is contributed by the electromagnetic self-force, and implies that the external force is *smaller* for a charged particle. This means that the electromagnetic self-force acting on the particle is *directed outward* and given by

$$f^r_{\rm self} = \frac{e^2 M}{r_0^3} f^{1/2}. \tag{1.50}$$

This is a *repulsive force*. It was shown by Zel'nikov and Frolov [22] that the same expression applies to a static charge outside a Reissner-Nordström black hole of mass M and charge Q, provided that f is replaced by the more general expression $f = 1 - 2M/r_0 + Q^2/r_0^2$.

The repulsive nature of the electromagnetic self-force acting on a static charge outside a black hole is unexpected. In an attempt to gain some intuition about this result, it is useful to recall that a black-hole horizon always acts as perfect conductor, because the electrostatic potential $\varphi := -A_t$ is necessarily uniform across its surface. It is then tempting to imagine that the self-force should result from a fictitious distribution of induced charge on the horizon, and that it could be estimated on the basis of an elementary model involving a spherical conductor. Let us, therefore, calculate the electric field produced by a point charge e situated outside a spherical conductor of radius R. The charge is placed at a distance r_0 from the centre of the conductor, which is taken at first to be grounded. The electrostatic potential produced by the charge can easily be obtained with the method of images. It is found that an image charge $e' = -eR/r_0$ is situated at a distance $r_0' = R^2/r_0$ from the centre of the conductor, and the potential is given by $\varphi = e/s + e'/s'$, where s is the distance to the charge, while s' is the distance to the image charge. The first term can be identified with the singular potential $\varphi_{\rm S}$, and the associated electric field exerts no force on the point charge. The second term is the regular potential $\varphi_{\rm R}$, and the associated field is entirely responsible for the self-force. The regular electric field is $E^r_{\rm R} = -\partial_r \varphi_{\rm R}$, and the self-force is $f^r_{\rm self} = eE^r_{\rm R}$. A simple computation returns

$$f^r_{\rm self} = -\frac{e^2 R}{r_0^3(1 - R^2/r_0^2)}. \tag{1.51}$$

This is an attractive self-force, because the total induced charge on the conducting surface is equal to e', which is opposite in sign to e. With R identified with M up to a numerical factor, we find that our intuition has produced the expected factor of $e^2 M/r_0^3$, but that it gives rise to the *wrong sign* for the self-force. An attempt to refine this computation by removing the net charge e' on the conductor (to mimic more closely the black-hole horizon, which cannot support a net charge) produces a wrong dependence on r_0 in addition to the same wrong sign. In this case the conductor is maintained at a constant potential $\phi_0 = -e'/R$, and the situation involves a second image charge $-e'$ situated at $r = 0$. It is easy to see that in this case,

$$f^r_{\rm self} = -\frac{e^2 R^3}{r_0^5(1 - R^2/r_0^2)}. \tag{1.52}$$

This is still an attractive force, which is weaker than the force of Eq. (1.51) by a factor of $(R/r_0)^2$; the force is now exerted by an image dipole instead of a single image charge.

The computation of the self-force in the black-hole case is almost as straightforward. The exact solution to Maxwell's equations that describes a point charge e situated $r = r_0$ and $\theta = 0$ in the Schwarzschild spacetime is given by

$$\varphi = \varphi^{\rm S} + \varphi^{\rm R}, \tag{1.53}$$

where

$$\varphi^{\rm S} = \frac{e}{r_0 r} \frac{(r - M)(r_0 - M) - M^2\cos\theta}{\bigl[(r-M)^2 - 2(r-M)(r_0-M)\cos\theta + (r_0-M)^2 - M^2\sin^2\theta\bigr]^{1/2}}, \tag{1.54}$$

is the solution first discovered by Copson in 1928 [23], while

$$\varphi^{\rm R} = \frac{eM/r_0}{r} \tag{1.55}$$

is the monopole field that was added by Linet [24] to obtain the correct asymptotic behaviour $\varphi \sim e/r$ when r is much larger than r_0. It is easy to see that Copson's potential behaves as $e(1 - M/r_0)/r$ at large distances, which reveals that in addition to e, $\varphi^{\rm S}$ comes with an additional (and unphysical) charge $-eM/r_0$ situated at $r = 0$. This charge must be removed by adding to $\varphi^{\rm S}$ a potential that (i) is a solution to the vacuum Maxwell equations, (ii) is regular everywhere except at $r = 0$, and (iii) carries the opposite charge $+eM/r_0$; this potential must be a pure monopole, because higher multipoles would produce a singularity on the horizon, and it is given uniquely by $\varphi^{\rm R}$. The Copson solution was generalized to Reissner-Nordström spacetime by Léauté and Linet [25], who also showed that the regular potential of Eq. (1.55) requires no modification.

The identification of Copson's potential with the singular potential $\varphi^{\rm S}$ is motivated by the fact that its associated electric field $F^{\rm S}_{tr} = \partial_r \varphi^{\rm S}$ is isotropic around the charge e and therefore exerts no force. The

self-force comes entirely from the monopole potential, which describes a (fictitious) charge $+eM/r_0$ situated at $r = 0$. Because this charge is of the same sign as the original charge e, the self-force is repulsive. More precisely stated, we find that the regular piece of the electric field is given by

$$F^{\mathrm{R}}_{tr} = -\frac{eM/r_0}{r^2}, \tag{1.56}$$

and that it produces the self-force of Eq. (1.50). The simple picture described here, in which the electromagnetic self-force is produced by a fictitious charge eM/r_0 situated at the centre of the black hole, is not easily extracted from the derivation presented originally by Smith and Will [21]. To the best of our knowledge, the monopolar origin of the self-force was first noticed by Alan Wiseman [26]. (In his paper, Wiseman computed the *scalar self-force* acting on a static particle in Schwarzschild spacetime, and found a zero answer. In this case, the analogue of the Copson solution for the scalar potential happens to satisfy the correct asymptotic conditions, and there is no need to add another solution to it. Because the scalar potential is precisely equal to the singular potential, the self-force vanishes.)

We should remark that the identification of φ^S and φ^R with the Detweiler-Whiting singular and regular fields, respectively, is a matter of conjecture. Although φ^S and φ^R satisfy the essential properties of the Detweiler-Whiting decomposition — being, respectively, a regular homogenous solution and a singular solution sourced by the particle — one should accept the possibility that they may not be the actual Detweiler-Whiting fields. It is a topic for future research to investigate the precise relation between the Copson field and the Detweiler-Whiting singular field.

It is instructive to compare the electromagnetic self-force produced by the presence of a grounded conductor to the self-force produced by the presence of a black hole. In the case of a conductor, the total induced charge on the conducting surface is $e' = -eR/r_0$, and it is this charge that is responsible for the attractive self-force; the induced charge is supplied by the electrodes that keep the conductor grounded. In the case of a black hole, there is no external apparatus that can supply such a charge, and the total induced charge on the horizon necessarily vanishes. The origin of the self-force is therefore very different in this case. As we have seen, the self-force is produced by a fictitious charge eM/r_0 situated at the centre of black hole; and because this charge is positive, the self-force is repulsive.

1.11 Organization of this review

After a detailed review of the literature in Sec. 2, the main body of the review begins in Part I (Secs. 3 to 7) with a description of the general theory of bitensors, the name designating tensorial functions of two points in spacetime. We introduce Synge's world function $\sigma(x, x')$ and its derivatives in Sec. 3, the parallel propagator $g^{\alpha}{}_{\alpha'}(x, x')$ in Sec. 5, and the van Vleck determinant $\Delta(x, x')$ in Sec. 7. An important portion of the theory (covered in Secs. 4 and 6) is concerned with the expansion of bitensors when x is very close to x'; expansions such as those displayed in Eqs. (1.23) and (1.24) are based on these techniques. The presentation in Part I borrows heavily from Synge's book [14] and the article by DeWitt and Brehme [4]. These two sources use different conventions for the Riemann tensor, and we have adopted Synge's conventions (which agree with those of Misner, Thorne, and Wheeler [27]). The reader is therefore warned that formulae derived in Part I may look superficially different from those found in DeWitt and Brehme.

In Part II (Secs. 8 to 11) we introduce a number of coordinate systems that play an important role in later parts of the review. As a warmup exercise we first construct (in Sec. 8) Riemann normal coordinates in a neighbourhood of a reference point x'. We then move on (in Sec. 9) to Fermi normal coordinates [28], which are defined in a neighbourhood of a world line γ. The retarded coordinates, which are also based at a world line and which were briefly introduced in Sec. 1.5, are covered systematically in Sec. 10. The relationship between Fermi and retarded coordinates is worked out in Sec. 11, which also locates the advanced point $z(v)$ associated with a field point x. The presentation in Part II borrows heavily from Synge's book [14]. In fact, we are much indebted to Synge for initiating the construction of retarded coordinates in a neighbourhood of a world line. We have implemented his program quite differently (Synge was interested in a large neighbourhood of the world line in a weakly curved spacetime, while we are interested in a small neighbourhood in a strongly curved spacetime), but the idea is originally his.

In Part III (Secs. 12 to 16) we review the theory of Green's functions for (scalar, vectorial, and tensorial) wave equations in curved spacetime. We begin in Sec. 12 with a pedagogical introduction to the retarded

and advanced Green's functions for a massive scalar field in flat spacetime; in this simple context the all-important Hadamard decomposition [29] of the Green's function into "light-cone" and "tail" parts can be displayed explicitly. The invariant Dirac functional is defined in Sec. 13 along with its restrictions on the past and future null cones of a reference point x'. The retarded, advanced, singular, and regular Green's functions for the scalar wave equation are introduced in Sec. 14. In Secs. 15 and 16 we cover the vectorial and tensorial wave equations, respectively. The presentation in Part III is based partly on the paper by DeWitt and Brehme [4], but it is inspired mostly by Friedlander's book [30]. The reader should be warned that in one important aspect, our notation differs from the notation of DeWitt and Brehme: While they denote the tail part of the Green's function by $-v(x,x')$, we have taken the liberty of eliminating the silly minus sign and call it instead $+V(x,x')$. The reader should also note that all our Green's functions are normalized in the same way, with a factor of -4π multiplying a four-dimensional Dirac functional of the right-hand side of the wave equation. (The gravitational Green's function is sometimes normalized with a -16π on the right-hand side.)

In Part IV (Secs. 17 to 19) we compute the retarded, singular, and regular fields associated with a point scalar charge (Sec. 17), a point electric charge (Sec. 18), and a point mass (Sec. 19). We provide two different derivations for each of the equations of motion. The first type of derivation was outlined previously: We follow Detweiler and Whiting [13] and postulate that only the regular field exerts a force on the particle. In the second type of derivation we take guidance from Quinn and Wald [7] and postulate that the net force exerted on a point particle is given by an average of the retarded field over a surface of constant proper distance orthogonal to the world line — this rest-frame average is easily carried out in Fermi normal coordinates. The averaged field is still infinite on the world line, but the divergence points in the direction of the acceleration vector and it can thus be removed by mass renormalization. Such calculations show that while the singular field does not affect the motion of the particle, it nonetheless contributes to its inertia.

In Part V (Secs. 20 to 23), we show that at linear order in the body's mass m, an extended body behaves just as a point mass, and except for the effects of the body's spin, the world line representing its mean motion is governed by the MiSaTaQuWa equation. At this order, therefore, the picture of a point particle interacting with its own field, and the results obtained from this picture, is justified. Our derivation utilizes the method of matched asymptotic expansions, with an inner expansion accurate near the body and an outer expansion accurate everywhere else. The equation of motion of the body's world line, suitably defined, is calculated by solving the Einstein equation in a buffer region around the body, where both expansions are accurate.

Concluding remarks are presented in Sec. 24, and technical developments that are required in Part V are relegated to Appendices. Throughout this review we use geometrized units and adopt the notations and conventions of Misner, Thorne, and Wheeler [27].

1.12 Changes relative to the 2004 edition

This 2010 version of the review is a major update of the original article published in 2004. Two additional authors, Adam Pound and Ian Vega, have joined the article's original author, and each one has contributed a major piece of the update. The literature survey presented in Sec. 2 was contributed by Ian Vega, and Part V (Secs. 20 to 23) was contributed by Adam Pound. Part V replaces a section of the 2004 article in which the motion of a small black hole was derived by the method of matched asymptotic expansions; this material can still be found in Ref. [15], but Pound's work provides a much more satisfactory foundation for the gravitational self-force. The case study of Sec. 1.10 is new, and the "exact" formulation of the dynamics of a point mass in Sec. 19.1 is a major improvement from the original article. The concluding remarks of Sec. 24, contributed mostly by Adam Pound, are also updated from the 2004 article.

Acknowledgments

Our understanding of the work presented in this review was shaped by a series of annual meetings named after the movie director Frank Capra. The first of these meetings took place in 1998 and was held at Capra's ranch in Southern California; the ranch now belongs to Caltech, Capra's alma mater. Subsequent meetings were held in Dublin (1999), Pasadena (2000), Potsdam (2001), State College PA (2002), Kyoto (2003), Brownsville (2004), Oxford (2005), Milwaukee (2006), Hunstville (2007), Orléans (2008), Bloomington (2009), and Wa-

terloo (2010). At these meetings and elsewhere we have enjoyed many instructive conversations with Paul Anderson, Warren Anderson, Patrick Brady, Claude Barrabès, Leor Barack, Luc Blanchet, Lior Burko, Manuella Campanelli, Marc Casals, Peter Diener, Steve Detweiler, Sam Dolan, Steve Drasco, Scott Field, Eanna Flanagan, John Friedman, Ryuichi Fujita, Chad Galley, Darren Golbourn, Sam Gralla, Roland Haas, Abraham Harte, Wataru Hikida, Tanja Hinderer, Scott Hughes, Werner Israel, Toby Keidl, Dong-hoon Kim, Alexandre Le Tiec, Carlos Lousto, Eirini Messaritaki, Yasushi Mino, Hiroyuki Nakano, Amos Ori, Larry Price, Eran Rosenthal, Norichika Sago, Misao Sasaki, Abhay Shah, Carlos Sopuerta, Alessandro Spallicci, Hideyuki Tagoshi, Takahiro Tanaka, Jonathan Thornburg, Bill Unruh, Bob Wald, Niels Warburton, Barry Wardell, Alan Wiseman, and Bernard Whiting. This work was supported by the Natural Sciences and Engineering Research Council of Canada.

2 Computing the self-force: a 2010 literature survey

Much progress has been achieved in the development of practical methods for computing the self-force. We briefly summarize these efforts in this section, with the goal of introducing the main ideas and some key issues. A more detailed coverage of the various implementations can be found in Barack's excellent review [31]. The 2005 collection of reviews published in *Classical and Quantum Gravity* [32] is also recommended for an introduction to the various aspects of self-force theory and numerics. Among our favourites in this collection are the reviews by Detweiler [33] and Whiting [34].

An important point to bear in mind is that all the methods covered here mainly compute the self-force on a particle moving on a *fixed world line* of the background spacetime. A few numerical codes based on the radiative approximation have allowed orbits to evolve according to energy and angular-momentum balance. As will be emphasized below, however, these calculations miss out on important conservative effects that are only accounted for by the full self-force. Work is currently underway to develop methods to let the self-force alter the motion of the particle in a self-consistent manner.

2.1 Early work: DeWitt and DeWitt; Smith and Will

The first evaluation of the electromagnetic self-force in curved spacetime was carried out by DeWitt and DeWitt [35] for a charge moving freely in a weakly curved spacetime characterized by a Newtonian potential $\Phi \ll 1$. In this context the right-hand side of Eq. (1.33) reduces to the tail integral, because the particle moves in a vacuum region of the spacetime, and there is no external force acting on the charge. They found that the spatial components of the self-force are given by

$$\boldsymbol{f}_{\rm em} = e^2 \frac{M}{r^3}\,\hat{\boldsymbol{r}} + \frac{2}{3} e^2 \frac{d\boldsymbol{g}}{dt}, \tag{2.1}$$

where M is the total mass contained in the spacetime, $r = |\boldsymbol{x}|$ is the distance from the centre of mass, $\hat{\boldsymbol{r}} = \boldsymbol{x}/r$, and $\boldsymbol{g} = -\boldsymbol{\nabla}\Phi$ is the Newtonian gravitational field. (In these expressions the bold-faced symbols represent vectors in three-dimensional flat space.) The first term on the right-hand side of Eq. (2.1) is a conservative correction to the Newtonian force $m\boldsymbol{g}$. The second term is the standard radiation-reaction force; although it comes from the tail integral, this is the same result that would be obtained in flat spacetime if an external force $m\boldsymbol{g}$ were acting on the particle. This agreement is necessary, but remarkable!

A similar expression was obtained by Pfenning and Poisson [36] for the case of a scalar charge. Here

$$\boldsymbol{f}_{\rm scalar} = 2\xi q^2 \frac{M}{r^3}\,\hat{\boldsymbol{r}} + \frac{1}{3} q^2 \frac{d\boldsymbol{g}}{dt}, \tag{2.2}$$

where ξ is the coupling of the scalar field to the spacetime curvature; the conservative term disappears when the field is minimally coupled. Pfenning and Poisson also computed the gravitational self-force acting on a point mass moving in a weakly curved spacetime. The expression they obtained is in complete agreement (within its domain of validity) with the standard post-Newtonian equations of motion.

The force required to hold an electric charge in place in a Schwarzschild spacetime was computed, without approximations, by Smith and Will [21]. As we reviewed previously in Sec. 1.10, the self-force contribution to the total force is given by

$$f^r_{\rm self} = e^2 \frac{M}{r^3} f^{1/2}, \tag{2.3}$$

where M is the black-hole mass, r the position of the charge (in Schwarzschild coordinates), and $f := 1 - 2M/r$. When $r \gg M$, this expression agrees with the conservative term in Eq. (2.1). This result was generalized to Reissner-Nordström spacetime by Zel'nikov and Frolov [22]. Wiseman [26] calculated the self-force acting on a static scalar charge in Schwarzschild spacetime. He found that in this case the self-force vanishes. This result is not incompatible with Eq. (2.2), even for nonminimal coupling, because the computation of the weak-field self-force requires the presence of matter, while Wiseman's scalar charge lives in a purely vacuum spacetime.

2.2 Mode-sum method

Self-force calculations involving a sum over modes were pioneered by Barack and Ori [37,38], and the method was further developed by Barack, Ori, Mino, Nakano, and Sasaki [39–44]; a somewhat related approach was also considered by Lousto [45]. It has now emerged as the method of choice for self-force calculations in spacetimes such as Schwarzschild and Kerr. Our understanding of the method was greatly improved by the Detweiler-Whiting decomposition [13] of the retarded field into singular and regular pieces, as outlined in Secs. 1.4 and 1.8, and subsequent work by Detweiler, Whiting, and their collaborators [46].

Detweiler-Whiting decomposition; mode decomposition; regularization parameters

For simplicity we consider the problem of computing the self-force acting on a particle with a scalar charge q moving on a world line γ. (The electromagnetic and gravitational problems are conceptually similar, and they will be discussed below.) The potential Φ produced by the particle satisfies Eq. (1.34), which we rewrite schematically as

$$\Box\Phi = q\delta(x,z), \tag{2.4}$$

where $\Box$ is the wave operator in curved spacetime, and $\delta(x,z)$ represents a distributional source that depends on the world line γ through its coordinate representation $z(\tau)$. From the perspective of the Detweiler-Whiting decomposition, the scalar self-force is given by

$$F_\alpha = q\nabla_\alpha\Phi_{\rm R} := q\bigl(\nabla_\alpha\Phi - \nabla_\alpha\Phi_{\rm S}\bigr), \tag{2.5}$$

where Φ, $\Phi_{\rm S}$, and $\Phi_{\rm R}$ are the retarded, singular, and regular potentials, respectively. To evaluate the self-force, then, is to compute the gradient of the regular potential.

From the point of view of Eq (2.5), the task of computing the self-force appears conceptually straightforward: Either (i) compute the retarded and singular potentials, subtract them, and take a gradient of the difference; or (ii) compute the gradients of the retarded and singular potentials, and then subtract the gradients. Indeed, this is the basic idea for most methods of self-force computations. However, the apparent simplicity of this sequence of steps is complicated by the following facts: (i) except for a very limited number of cases, the retarded potential of a point particle cannot be computed analytically and must therefore be obtained by numerical means; and (ii) both the retarded and singular potential diverge at the particle's position. Thus, any sort of subtraction will generally have to be performed numerically, and for this to be possible, one requires representations of the retarded and singular potentials (and/or their gradients) in terms of finite quantities.

In a mode-sum method, these difficulties are overcome with a decomposition of the potential in spherical-harmonic functions:

$$\Phi = \sum_{lm} \Phi^{lm}(t,r) Y^{lm}(\theta,\phi). \tag{2.6}$$

When the background spacetime is spherically symmetric, Eq. (2.4) gives rise to a fully decoupled set of reduced wave equations for the mode coefficients $\Phi^{lm}(t,r)$, and these are easily integrated with simple numerical methods. The dimensional reduction of the wave equation implies that each $\Phi^{lm}(t,r)$ is finite and continuous (though nondifferentiable) at the position of the particle. There is, therefore, no obstacle to evaluating each l-mode of the field, defined by

$$(\nabla_\alpha\Phi)_l := \lim_{x\to z} \sum_{m=-l}^{l} \nabla_\alpha[\Phi^{lm}(t,r)Y^{lm}(\theta,\phi)]. \tag{2.7}$$

The sum over modes, however, must reproduce the singular field evaluated at the particle's position, and this is infinite; the mode sum, therefore, does not converge.

Fortunately, there is a piece of each l-mode that does not contribute to the self-force, and that can be subtracted out; this piece is the corresponding l-mode of the singular field $\nabla_\alpha \Phi_{\rm S}$. Because the retarded and singular fields share the same singularity structure near the particle's world line (as described in Sec. 1.6), the subtraction produces a mode decomposition of the regular field $\nabla_\alpha \Phi_{\rm R}$. And because this field is regular at the particle's position, the sum over all modes $q(\nabla_\alpha \Phi_{\rm R})_l$ is guaranteed to converge to the correct value for the self-force. The key to the mode-sum method, therefore, is the ability to express the singular field as a mode decomposition.

This can be done because the singular field, unlike the retarded field, can always be expressed as a local expansion in powers of the distance to the particle; such an expansion was displayed in Eqs. (1.28) and (1.29). (In a few special cases the singular field is actually known exactly [23, 24, 47–49].) This local expansion can then be turned into a multipole decomposition. Barack and Ori [31, 39, 41–43], and then Mino, Nakano, and Sasaki [44], were the first to show that this produces the following generic structure:

$$(\nabla_\alpha \Phi_{\rm S})_l = (l+\tfrac{1}{2})A_\alpha + B_\alpha + \frac{C_\alpha}{l+\frac{1}{2}} + \frac{D_\alpha}{(l-\frac{1}{2})(l+\frac{3}{2})} + \frac{E_\alpha}{(l-\frac{3}{2})(l-\frac{1}{2})(l+\frac{3}{2})(l+\frac{5}{2})} + \cdots, \tag{2.8}$$

where A_α, B_α, C_α, and so on are l-independent functions that depend on the choice of field (i.e. scalar, electromagnetic, or gravitational), the choice of spacetime, and the particle's state of motion. These so-called *regularization parameters* are now ubiquitous in the self-force literature, and they can all be determined from the local expansion for the singular field. The number of regularization parameters that can be obtained depends on the accuracy of the expansion. For example, expansions accurate through order r^0 such as Eqs. (1.28) and (1.29) permit the determination of A_α, B_α, and C_α; to obtain D_α one requires the terms of order r, and to get E_α the expansion must be carried out through order r^2. The particular polynomials in l that accompany the regularization parameters were first identified by Detweiler and his collaborators [46]. Because the D_α term is generated by terms of order r in the local expansion of the singular field, the sum of $[(l-\frac{1}{2})(l+\frac{3}{2})]^{-1}$ from $l=0$ to $l=\infty$ evaluates to zero. The sum of the polynomial in front of E_α also evaluates to zero, and this property is shared by all remaining terms in Eq. (2.8).

Mode sum

With these elements in place, the self-force is finally computed by implementing the mode-sum formula

$$\begin{aligned} F_\alpha &= q\sum_{l=0}^{L}\biggl[(\nabla_\alpha\Phi)_l - (l+\tfrac{1}{2})A_\alpha - B_\alpha - \frac{C_\alpha}{l+\frac{1}{2}} - \frac{D_\alpha}{(l-\frac{1}{2})(l+\frac{3}{2})} \\ &\quad - \frac{E_\alpha}{(l-\frac{3}{2})(l-\frac{1}{2})(l+\frac{3}{2})(l+\frac{5}{2})} - \cdots\biggr] + \text{remainder}, \end{aligned} \tag{2.9}$$

where the infinite sum over l is truncated to a maximum mode number L. (This truncation is necessary in practice, because in general the modes must be determined numerically.) The remainder consists of the remaining terms in the sum, from $l = L+1$ to $l=\infty$; it is easy to see that since the next regularization term would scale as l^{-6} for large l, the remainder scales as L^{-5}, and can be made negligible by summing to a suitably large value of l. This observation motivates the inclusion of the D_α and E_α terms within the mode sum, even though their complete sums evaluate to zero. These terms are useful because the sum must necessarily be truncated, and they permit a more rapid convergence of the mode sum. For example, exclusion of the D_α and E_α terms in Eq. (2.9) would produce a remainder that scales as L^{-1} instead of L^{-5}; while this is sufficient for convergence, the rate of convergence is too slow to permit high-accuracy computations. Rapid convergence therefore relies on a knowledge of as many regularization parameters as possible, but unfortunately these parameters are not easy to calculate. To date, only A_α, B_α, C_α, and D_α have been calculated for general orbits in Schwarzschild spacetime [46, 50], and only A_α, B_α, C_α have been calculated for orbits in Kerr spacetime [43]. It is possible, however, to estimate a few additional regularization parameters by fitting numerical results to the structure of Eq. (2.8); this clever trick was first exploited by Detweiler and his collaborators [46] to achieve extremely high numerical accuracies. This trick is now applied routinely in mode-sum computations of the self-force.

Case study: static electric charge in extreme Reissner-Nordström spacetime

The practical use of the mode-sum method can be illustrated with the help of a specific example that can be worked out fully and exactly. We consider, as in Sec. 1.10, an electric charge e held in place at position $r = r_0$ in the spacetime of an extreme Reissner-Nordström black hole of mass M and charge $Q = M$. The reason for selecting this spacetime resides in the resulting simplicity of the spherical-harmonic modes for the electromagnetic field.

The metric of the extreme Reissner-Nordström spacetime is given by

$$ds^2 = -f\,dt^2 + f^{-1}dr^2 + r^2 d\Omega^2, \tag{2.10}$$

where $f = (1 - M/r)^2$. The only nonzero component of the electromagnetic field tensor is $F_{tr} = -E_r$, and this is decomposed as

$$F_{tr} = \sum_{lm} F^{lm}_{tr}(r) Y^{lm}(\theta, \phi). \tag{2.11}$$

This field diverges at $r = r_0$, but the modes $F^{lm}_{tr}(r)$ are finite, though discontinuous. The multipole coefficients of the field are defined to be

$$(F_{tr})_l = \lim \sum_{m=-l}^{l} F^{lm}_{tr} Y^{lm}, \tag{2.12}$$

where the limit is taken in the direction of the particle's position. The charge can be placed on the axis $\theta = 0$, and this choice produces an axisymmetric field with contributions from $m = 0$ only. Because $Y^{l0} = [(2l+1)/4\pi]^{1/2} P_l(\cos\theta)$ and $P_l(1) = 1$, we have

$$(F_{tr})_l = \sqrt{\frac{2l+1}{4\pi}} \lim_{\Delta\to 0} F^{l0}_{tr}(r_0 + \Delta). \tag{2.13}$$

The sign of Δ is arbitrary, and $(F_{tr})_l$ depends on the direction in which r_0 is approached.

The charge density of a static particle can also be decomposed in spherical harmonics, and the mode coefficients are given by

$$r^2 j^{l0}_t = e\sqrt{\frac{2l+1}{4\pi}} f_0 \delta(r - r_0), \tag{2.14}$$

where $f_0 = (1 - M/r_0)^2$. If we let

$$\Phi^l := -r^2 F^{l0}_{tr}, \tag{2.15}$$

then Gauss's law in the extreme Reissner-Nordström spacetime can be shown to reduce to

$$(f\Phi')' - \frac{l(l+1)}{r^2}\Phi = 4\pi e\sqrt{\frac{2l+1}{4\pi}} f_0 \delta'(r - r_0), \tag{2.16}$$

in which a prime indicates differentiation with respect to r, and the index l on Φ is omitted to simplify the expressions. The solution to Eq. (2.16) can be expressed as $\Phi(r) = \Phi_>(r)\Theta(r - r_0) + \Phi_<(r)\Theta(r_0 - r)$, where $\Phi_>$ and $\Phi_<$ are each required to satisfy the homogeneous equation $(f\Phi')' - l(l+1)\Phi/r^2 = 0$, as well as the junction conditions

$$[\Phi] = 4\pi e\sqrt{\frac{2l+1}{4\pi}}, \qquad [\Phi'] = 0, \tag{2.17}$$

with $[\Phi] := \Phi_>(r_0) - \Phi_<(r_0)$ denoting the jump across $r = r_0$.

For $l = 0$ the general solution to the homogeneous equation is $c_1 r^* + c_2$, where c_1 and c_2 are constants and $r^* = \int f^{-1}\,dr$. The solution for $r < r_0$ must be regular at $r = M$, and we select $\Phi_<$ = constant. The solution for $r > r_0$ must produce a field that decays as r^{-2} at large r, and we again select $\Phi_>$ = constant. Since each constant is proportional to the total charge enclosed within a sphere of radius r, we arrive at

$$\Phi_< = 0, \qquad \Phi_> = \sqrt{4\pi}e, \qquad (l = 0). \tag{2.18}$$

For $l \neq 0$ the solutions to the homogeneous equation are

$$\Phi_< = c_1 e \left(\frac{r - M}{r_0 - M} \right)^l (lr + M) \tag{2.19}$$

and

$$\Phi_> = c_2 e \left(\frac{r_0 - M}{r - M} \right)^{l+1} \big[(l+1)r - M \big]. \tag{2.20}$$

The constants c_1 and c_2 are determined by the junction conditions, and we get

$$c_1 = -\sqrt{\frac{4\pi}{2l+1}} \frac{1}{r_0}, \qquad c_2 = \sqrt{\frac{4\pi}{2l+1}} \frac{1}{r_0}. \tag{2.21}$$

The modes of the electromagnetic field are now completely determined.

According to the foregoing results, and recalling the definition of Eq. (2.13), the multipole coefficients of the electromagnetic field at $r = r_0 + 0^+$ are given by

$$\big(F^>_{tr}\big)_0 = -\frac{e}{r_0^2}, \qquad \big(F^>_{tr}\big)_l = e\big(l + \tfrac{1}{2}\big)\left(-\frac{1}{r_0^2}\right) - \frac{e}{2r_0^3}(r_0 - 2M). \tag{2.22}$$

For $r = r_0 + 0^-$ we have instead

$$\big(F^<_{tr}\big)_0 = 0, \qquad \big(F^<_{tr}\big)_l = e\big(l + \tfrac{1}{2}\big)\left(+\frac{1}{r_0^2}\right) - \frac{e}{2r_0^3}(r_0 - 2M). \tag{2.23}$$

We observe that the multipole coefficients lead to a diverging mode sum. We also observe, however, that the multipole structure is identical to the decomposition of the singular field displayed in Eq. (2.8). Comparison of the two expressions allows us to determine the regularization parameters for the given situation, and we obtain

$$A = \mp \frac{e}{r_0^2}, \qquad B = -\frac{e}{2r_0^3}(r_0 - 2M), \qquad C = D = E = \cdots = 0. \tag{2.24}$$

Regularization of the mode sum via Eq. (2.9) reveals that the modes $l \neq 0$ give rise to the singular field, while the regular field comes entirely from the mode $l = 0$. In this case, therefore, we can state that the *exact expression* for the regular field evaluated at the position of the particle is $F^{\mathsf{R}}_{tr} = (F_{tr})_0 - \frac{1}{2}A - B$, or $F^{\mathsf{R}}_{tr}(r_0) = -eM/r_0^3$. Because the regular field must be a solution to the vacuum Maxwell equations, its monopole structure guarantees that its value at any position is given by

$$F^{\mathrm{R}}_{tr}(r) = -\frac{eM/r_0}{r^2}. \tag{2.25}$$

This is the field of an image charge $e' = +eM/r_0$ situated at the centre of the black hole.

The self-force acting on the static charge is then

$$f^r = -e\sqrt{f_0} F^{\mathrm{R}}_{tr}(r_0) = \frac{e^2 M}{r_0^3}\sqrt{f_0} = \frac{e^2 M}{r_0^3}(1 - M/r_0). \tag{2.26}$$

This expression agrees with the Smith-Will force of Eq. (1.50). The interpretation of the result in terms of an interaction between e and the image charge e' was elaborated in Sec. 1.10.

Computations in Schwarzschild spacetime

The mode-sum method was successfully implemented in Schwarzschild spacetime to compute the scalar and electromagnetic self-forces on a static particle [51, 52] . It was used to calculate the scalar self-force on a particle moving on a radial trajectory [53], circular orbit [46, 50, 54, 55], and a generic bound orbit [56]. It was also developed to compute the electromagnetic self-force on a particle moving on a generic bound orbit [57], as well as the gravitational self-force on a point mass moving on circular [58, 59] and eccentric orbits [60]. The mode-sum method was also used to compute unambiguous physical effects associated with the

gravitational self-force [61–63], and these results were involved in detailed comparisons with post-Newtonian theory [61, 63–66]. These achievements will be described in more detail in Sec. 2.6.

An issue that arises in computations of the electromagnetic and gravitational self-forces is the choice of gauge. While the self-force formalism is solidly grounded in the Lorenz gauge (which allows the formulation of a wave equation for the potentials, the decomposition of the retarded field into singular and regular pieces, and the computation of regularization parameters), it is often convenient to carry out the numerical computations in other gauges, such as the popular Regge-Wheeler gauge and the Chrzanowski radiation gauge described below. Compatibility of calculations carried out in different gauges has been debated in the literature. It is clear that the singular field is gauge invariant when the transformation between the Lorenz gauge and the adopted gauge is smooth on the particle's world line; in such cases the regularization parameters also are gauge invariant [19], the transformation affects the regular field only, and the self-force changes according to Eq. (1.49). The transformations between the Lorenz gauge and the Regge-Wheeler and radiation gauges are not regular on the world line, however, and in such cases the regularization of the retarded field must be handled with extreme care.

Computations in Kerr spacetime; metric reconstruction

The reliance of the mode-sum method on a spherical-harmonic decomposition makes it generally impractical to apply to self-force computations in Kerr spacetime. Wave equations in this spacetime are better analyzed in terms of a *spheroidal*-harmonic decomposition, which simultaneously requires a Fourier decomposition of the field's time dependence. (The eigenvalue equation for the angular functions depends on the mode's frequency.) For a static particle, however, the situation simplifies, and Burko and Liu [67] were able to apply the method to calculate the self-force on a static scalar charge in Kerr spacetime.

More recently, Warburton and Barack [68] carried out mode-sum calculations of the scalar self-force on a particle moving on equatorial orbits of a Kerr black hole. They first solve for the spheroidal multipoles of the retarded potential, and then re-express them in terms of spherical-harmonic multipoles. Fortunately, they find that a spheroidal multipole is well represented by summing over a limited number of spherical multipoles. The Warburton-Barack work represents the first successful computations of the self-force in Kerr spacetime, and it reveals the interesting effect of the black hole's spin on the behaviour of the self-force.

The analysis of the scalar wave equation in terms of spheroidal functions and a Fourier decomposition permits a complete separation of the variables. For decoupling and separation to occur in the case of a gravitational perturbation, it is necessary to formulate the perturbation equations in terms of Newman-Penrose (NP) quantities [69], and to work with the Teukolsky equation that governs their behaviour. Several computer codes are now available that are capable of integrating the Teukolsky equation when the source is a point mass moving on an arbitrary geodesic of the Kerr spacetime. (A survey of these codes is given below.) Once a solution to the Teukolsky equation is at hand, however, there still remains the additional task of recovering the metric perturbation from this solution, a problem referred to as *metric reconstruction*.

Reconstruction of the metric perturbation from solutions to the Teukolsky equation was tackled in the past in the pioneering efforts of Chrzanowski [70], Cohen and Kegeles [71, 72], Stewart [73], and Wald [74]. These works have established a procedure, typically attributed to Chrzanowski, that returns the metric perturbation in a so-called radiation gauge. An important limitation of this method, however, is that it applies only to vacuum solutions to the Teukolsky equation. This makes the standard Chrzanowski procedure inapplicable in the self-force context, because a point particle must necessarily act as a source of the perturbation. Some methods were devised to extend the Chrzanowski procedure to accomodate point sources in specific circumstances [75, 76], but these were not developed sufficiently to permit the computation of a self-force. See Ref. [77] for a review of metric reconstruction from the perspective of self-force calculations.

A remarkable breakthrough in the application of metric-reconstruction methods in self-force calculations was achieved by Keidl, Wiseman, and Friedman [78–80], who were able to compute a self-force starting from a Teukolsky equation sourced by a point particle. They did it first for the case of an electric charge and a point mass held at a fixed position in a Schwarzschild spacetime [78], and then for the case of a point mass moving on a circular orbit around a Schwarzschild black hole [80]. The key conceptual advance is the realization that, according to the Detweiler-Whiting perspective, the self-force is produced by a regularized field that satisfies vacuum field equations in a neighbourhood of the particle. The regular field can therefore be submitted to the Chrzanowski procedure and reconstructed from a source-free solution to the Teukolsky

equation.

More concretely, suppose that we have access to the Weyl scalar ψ_0 produced by a point mass moving on a geodesic of a Kerr spacetime. To compute the self-force from this, one first calculates the singular Weyl scalar $\psi_0^{\rm S}$ from the Detweiler-Whiting singular field $h^{\rm S}_{\alpha\beta}$, and subtracts it from ψ_0. The result is a regularized Weyl scalar $\psi_0^{\rm R}$, which is a solution to the homogeneous Teukolsky equation. This sets the stage for the metric-reconstruction procedure, which returns (a piece of) the regular field $h^{\rm R}_{\alpha\beta}$ in the radiation gauge selected by Chrzanowski. The computation must be completed by adding the pieces of the metric perturbation that are not contained in ψ_0; these are referred to either as the nonradiative degrees of freedom (since ψ_0 is purely radiative), or as the $l = 0$ and $l = 1$ field multipoles (because the sum over multipoles that make up ψ_0 begins at $l = 2$). A method to complete the Chrzanowski reconstruction of $h^{\rm R}_{\alpha\beta}$ was devised by Keidl *et al.* [78, 80], and the end result leads directly to the gravitational self-force. The relevance of the $l = 0$ and $l = 1$ modes to the gravitational self-force was emphasized by Detweiler and Poisson [81].

Time-domain versus frequency-domain methods

When calculating the spherical-harmonic components $\Phi^{lm}(t, r)$ of the retarded potential Φ — refer back to Eq. (2.6) — one can choose to work either directly in the time domain, or perform a Fourier decomposition of the time dependence and work instead in the frequency domain. While the time-domain method requires the integration of a partial differential equation in t and r, the frequency-domain method gives rise to set of ordinary differential equations in r, one for each frequency ω. For particles moving on circular or slightly eccentric orbits in Schwarzschild spacetime, the frequency spectrum is limited to a small number of discrete frequencies, and a frequency-domain method is easy to implement and yields highly accurate results. As the orbital eccentricity increases, however, the frequency spectrum broadens, and the computational burden of summing over all frequency components becomes more significant. Frequency-domain methods are less efficient for large eccentricities, the case of most relevance for extreme-mass-ratio inspirals, and it becomes advantageous to replace them with time-domain methods. (See Ref. [82] for a quantitative study of this claim.) This observation has motivated the development of accurate evolution codes for wave equations in $1 + 1$ dimensions.

Such codes must be able to accomodate point-particle sources, and various strategies have been pursued to represent a Dirac distribution on a numerical grid, including the use of very narrow Gaussian pulses [83–85] and of "finite impulse representations" [86]. These methods do a good job with waveform and radiative flux calculations far away from the particle, but are of very limited accuracy when computing the potential in a neighborhood of the particle. A numerical method designed to provide an *exact representation* of a Dirac distribution in a time-domain computation was devised by Lousto and Price [87] (see also Ref. [88]). It was implemented by Haas [56, 57] for the specific purpose of evaluating $\Phi^{lm}(t, r)$ at the position of particle and computing the self-force. Similar codes were developed by other workers for scalar [89] and gravitational [58, 60] self-force calculations.

Most extant time-domain codes are based on finite-difference techniques, but codes based on pseudo-spectral methods have also been developed [55, 90–92]. Spectral codes are a powerful alternative to finite-difference codes, especially when dealing with smooth functions, because they produce much faster convergence. The fact that self-force calculations deal with point sources and field modes that are not differentiable might suggest that spectral convergence should not be expected in this case. This objection can be countered, however, by placing the particle at the boundary between two spectral domains. Functions are then smooth in each domain, and discontinuities are handled by formulating appropriate boundary conditions; spectral convergence is thereby achieved.

2.3 Effective-source method

The mode-sum methods reviewed in the preceding subsection have been developed and applied extensively, but they do not exhaust the range of approaches that may be exploited to compute a self-force. Another set of methods, devised by Barack and his collaborators [93–95] as well as Vega and his collaborators [89, 96, 97], begin by recognizing that an approximation to the exact singular potential can be used to regularize the delta-function source term of the original field equation. We shall explain this idea in the simple context of a scalar potential Φ.

We continue to write the wave equation for the retarded potential Φ in the schematic form

$$\Box\Phi = q\delta(x,z), \tag{2.27}$$

where $\Box$ is the wave operator in curved spacetime, and $\delta(x,z)$ is a distributional source term that depends on the particle's world line γ through its coordinate representation $z(\tau)$. By construction, the exact singular potential $\Phi_{\rm S}$ satisfies the same equation, and an approximation to the singular potential, denoted $\tilde{\Phi}_{\rm S}$, will generally satisfy an equation of the form

$$\Box\tilde{\Phi}_{\rm S} = q\delta(x,x_0) + O(r^n) \tag{2.28}$$

for some integer $n > 0$, where r is a measure of distance to the world line. A "better" approximation to the singular potential is one with a higher value of n. From the approximated singular potential we form an approximation to the regular potential by writing

$$\tilde{\Phi}_{\rm R} := \Phi - W\tilde{\Phi}_{\rm S}, \tag{2.29}$$

where W is a window function whose properties will be specified below. The approximated regular potential is governed by the wave equation

$$\Box\tilde{\Phi}_{\rm R} = q\delta(x,z) - \Box\big(W\tilde{\Phi}_{\rm S}\big) := S(x,z), \tag{2.30}$$

and the right-hand side of this equation defines the effective source term $S(x,z)$. This equation is much less singular than Eq. (2.27), and it can be integrated using numerical methods designed to handle smooth functions.

To see this, we write the effective source more specifically as

$$S(x,z) = -\tilde{\Phi}_{\rm S}\Box W - 2\nabla_\alpha W\nabla^\alpha\tilde{\Phi}_{\rm S} - W\Box\tilde{\Phi}_{\rm S} + q\delta(x,z). \tag{2.31}$$

With the window function W designed to approach unity as $x \to z$, we find that the delta function that arises from the third term on the right-hand side precisely cancels out the fourth term. To keep the other terms in S well behaved on the world line, we further restrict the window function to satisfy $\nabla_\alpha W = O(r^p)$ with $p \geq 2$; this ensures that multiplication by $\nabla_\alpha\tilde{\Phi}_{\rm S} = O(r^{-2})$ leaves behind a bounded quantity. In addition, we demand that $\Box W = O(r^q)$ with $q \geq 1$, so that multiplication by $\tilde{\Phi}_{\rm S} = O(r^{-1})$ again produces a bounded quantity. It is also useful to require that $W(x)$ have compact (spatial) support, to ensure that the effective source term $S(x,z)$ does not extend beyond a reasonably small neighbourhood of the world line; this property also has the virtue of making $\tilde{\Phi}_{\rm R}$ precisely equal to the retarded potential Φ outside the support of the window function. This implies, in particular, that $\tilde{\Phi}_{\rm R}$ can be used directly to compute radiative fluxes at infinity. Another considerable virtue of these specifications for the window function is that they guarantee that the gradient of $\tilde{\Phi}_{\rm R}$ is directly tied to the self-force. We indeed see that

$$\begin{aligned}
\lim_{x\to z}\nabla_\alpha\tilde{\Phi}_{\rm R} &= \lim_{x\to z}\big(\nabla_\alpha\Phi - W\nabla_\alpha\tilde{\Phi}_{\rm S}\big) - \lim_{x\to z}\tilde{\Phi}_{\rm S}\nabla_\alpha W \\
&= \lim_{x\to z}\big(\nabla_\alpha\Phi - \nabla_\alpha\tilde{\Phi}_{\rm S}\big) \\
&= q^{-1}F_\alpha,
\end{aligned} \tag{2.32}$$

with the second line following by virtue of the imposed conditions on W, and the third line from the properties of the approximated singular field.

The effective-source method therefore consists of integrating the wave equation

$$\Box\tilde{\Phi}_{\rm R} = S(x,z), \tag{2.33}$$

for the approximated regular potential $\tilde{\Phi}_{\rm R}$, with a source term $S(x,z)$ that has become a regular function (of limited differentiability) of the spacetime coordinates x. The method is also known as a "puncture approach," in reference to a similar regularization strategy employed in numerical relativity. It is well suited to a $3+1$ integration of the wave equation, which can be implemented on mature codes already in circulation within the numerical-relativity community. An important advantage of a $3+1$ implementation is that it is largely

indifferent to the choice of background spacetime, and largely insensitive to the symmetries possessed by this spacetime; a self-force in Kerr spacetime is in principle just as easy to obtain as a self-force in Schwarzschild spacetime.

The method is also well suited to a self-consistent implementation of the self-force, in which the motion of the particle is not fixed in advance, but determined by the action of the computed self-force. This amounts to combining Eq. (2.33) with the self-force equation

$$m\frac{Du^\mu}{d\tau} = q\bigl(g^{\mu\nu} + u^\mu u^\nu\bigr)\nabla_\nu \tilde{\Phi}_{\rm R}, \tag{2.34}$$

in which the field is evaluated on the dynamically determined world line. The system of equations is integrated jointly, and self-consistently. The $3+1$ version of the effective-source approach presents a unique opportunity for the numerical-relativity community to get involved in self-force computations, with only a minimal amount of infrastructure development. This was advocated by Vega and Detweiler [89], who first demonstrated the viability of the approach with a $1+1$ time-domain code for a scalar charge on a circular orbit around a Schwarzschild black hole. An implementation with two separate $3+1$ codes imported from numerical relativity was also accomplished [96].

The work of Barack and collaborators [93, 94] is a particular implementation of the effective-source approach in a $2+1$ numerical calculation of the scalar self-force in Kerr spacetime. (See also the independent implementation by Lousto and Nakano [98].) Instead of starting with Eq. (2.27), they first decompose Φ according to

$$\Phi(x) = \sum_m \Phi^m(t,r,\theta)\exp(im\phi) \tag{2.35}$$

and formulate reduced wave equations for the Fourier coefficients Φ^m. Each coefficient is then regularized with an appropriate singular field $\tilde{\Phi}^m_{\rm S}$, which eliminates the delta-funtion from Eq. (2.27). This gives rise to regularized source terms for the reduced wave equations, which can then be integrated with a $2+1$ evolution code. In the final stage of the computation, the self-force is recovered by summing over the regularized Fourier coefficients. This strategy, known as the *m-mode regularization scheme*, is currently under active development. Recently it was successfully applied by Dolan and Barack [95] to compute the self-force on a scalar charge in circular orbit around a Schwarzschild black hole.

2.4 Quasilocal approach with “matched expansions”

As was seen in Eqs. (1.33), (1.40), and (1.47), the self-force can be expressed as an integral over the past world line of the particle, the integrand involving the Green’s function for the appropriate wave equation. Attempts have been made to compute the Green’s function directly [35, 36, 47, 48], and to evaluate the world-line integral. The quasilocal approach, first introduced by Anderson and his collaborators [99–102], is based on the expectation that the world-line integral might be dominated by the particle’s recent past, so that the Green’s function can be represented by its Hadamard expansion, which is restricted to the normal convex neighbourhood of the particle’s current position. To help with this enterprise, Ottewill and his collaborators [103–106] have pushed the Hadamard expansion to a very high order of accuracy, building on earlier work by Décanini and Folacci [107].

The weak-field calculations performed by DeWitt and DeWitt [35] and Pfenning and Poisson [36] suggest that the world-line integral is not, in fact, dominated by the recent past. Instead, most of the self-force is produced by signals that leave the particle at some time in the past, scatter off the central mass, and reconnect with the particle at the current time; such signals mark the boundary of the normal convex neighbourhood. The quasilocal evaluation of the world-line integral must therefore be supplemented with contributions from the distant past, and this requires a representation of the Green’s function that is not limited to the normal convex neighbourhood. In some spacetimes it is possible to express the Green’s function as an expansion in quasi-normal modes, as was demonstrated by Casals and his collaborators for a static scalar charge in the Nariai spacetime [108]. Their study provided significant insights into the geometrical structure of Green’s functions in curved spacetime, and increased our understanding of the non-local character of the self-force.

2.5 Adiabatic approximations

The accurate computation of long-term waveforms from extreme-mass-ratio inspirals (EMRIs) involves a lengthy sequence of calculations that include the calculation of the self-force. One can already imagine the difficulty of numerically integrating the coupled linearized Einstein equation for durations much longer than has ever been attempted by existing numerical codes. While doing so, the code would also have to evaluate the self-force reasonably often (if not at each time step) in order to remain close to the true dynamics of the point mass. Moreover, gravitational-wave data analysis via matched filtering require full evolutions of the sort just described for a large sample of systems parameters. All these considerations underlie the desire for simplified approximations to fully self-consistent self-force EMRI models.

The *adiabatic approximation* refers to one such class of potentially useful approximations. The basic assumption is that the secular effects of the self-force occur on a timescale that is much longer than the orbital period. In an extreme-mass-ratio binary, this assumption is valid during the early stage of inspiral; it breaks down in the final moments, when the orbit transitions to a quasi-radial infall called the plunge. From the adiabaticity assumption, numerous approximations have been formulated: For example, (i) since the particle's orbit deviates only slowly from geodesic motion, the self-force can be calculated from a field sourced by a geodesic; (ii) since the radiation-reaction timescale t_{rr}, over which a significant shrinking of the orbit occurs due to the self-force, is much longer than the orbital period, periodic effects of the self-force can be neglected; and (iii) conservative effects of the self-force can be neglected (the *radiative approximation*).

A seminal example of an adiabatic approximation is the Peters-Mathews formalism [109, 110], which determines the long-term evolution of a binary orbit by equating the time-averaged rate of change of the orbital energy E and angular momentum L to, respectively, the flux of gravitational-wave energy and angular momentum at infinity. This formalism was used to successfully predict the decreasing orbital period of the Hulse-Taylor pulsar, before more sophisticated methods, based on post-Newtonian equations of motion expanded to 2.5PN order, were incorporated in times-of-arrival formulae.

In the hope of achieving similar success in the context of the self-force, considerable work has been done to formulate a similar approximation for the case of an extreme-mass-ratio inspiral [111–121]. Bound geodesics in Kerr spacetime are specified by three constants of motion — the energy E, angular momentum L, and Carter constant C. If one could easily calculate the rates of change of these quantities, using a method analogous to the Peters-Mathews formalism, then one could determine an approximation to the long-term orbital evolution of the small body in an EMRI, avoiding the lengthy process of regularization involved in the direct integration of the self-forced equation of motion. In the early 1980s, Gal'tsov [122] showed that the average rates of change of E and L, as calculated from balance equations that assume geodesic source motion, agree with the averaged rates of change induced by a self-force constructed from a radiative Green's function defined as $G_{\rm rad} := \frac{1}{2}(G_- - G_+)$. As discussed in Sec. 1.4, this is equal to the regular two-point function $G_{\rm R}$ in flat spacetime, but $G_{\rm rad} \neq G_{\rm R}$ in curved spacetime; because of its time-asymmetry, it is purely dissipative. Mino [111], building on the work of Gal'tsov, was able to show that the true self-force and the dissipative force constructed from $G_{\rm rad}$ cause the same averaged rates of change of all three constants of motion, lending credence to the radiative approximation. Since then, the radiative Green's function was used to derive explicit expressions for the rates of change of E, L, and C in terms of the particle's orbit and wave amplitudes at infinity [117–119], and radiative approximations based on such expressions have been concretely implemented by Drasco, Hughes and their collaborators [115, 116, 123].

The relevance of the conservative part of the self-force — the part left out when using $G_{\rm rad}$ — was analyzed in numerous recent publications [121, 124–128]. As was shown by Pound *et al.* [125–127], neglect of the conservative effects of the self-force generically leads to long-term errors in the phase of an orbit and the gravitational wave it produces. These phasing errors are due to orbital precession and a direct shift in orbital frequency. This shift can be understood by considering a conservative force acting on a circular orbit: the force is radial, it alters the centripetal acceleration, and the frequency associated with a given orbital radius is affected. Despite these errors, a radiative approximation may still suffice for gravitational-wave detection [121]; for circular orbits, which have minimal conservative effects, radiative approximations may suffice even for parameter-estimation [128]. However, at this point in time, these analyses remain inconclusive because they all rely on extrapolations from post-Newtonian results for the conservative part of the self-force. For a more comprehensive discussion of these issues, the reader is referred to Ref. [121, 129].

Hinderer and Flanagan performed the most comprehensive study of these issues [130], utilizing a two-

timescale expansion [18, 131] of the field equations and self-forced equations of motion in an EMRI. In this method, all dynamical variables are written in terms of two time coordinates: a fast time t and a slow time $\tilde{t} := (m/M)t$. In the case of an EMRI, the dynamical variables are the metric and the phase-space variables of the world line. The fast-time dependence captures evolution on the orbital timescale $\sim M$, while the slow-time dependence captures evolution on the radiation-reaction timescale $\sim M^2/m$. One obtains a sequence of fast-time and slow-time equations by expanding the full equations in the limit of small m while treating the two time coordinates as independent. Solving the leading-order fast-time equation, in which $\tilde{t}$ is held fixed, yields a metric perturbation sourced by a geodesic, as one would expect from the linearized field equations for a point particle. The leading-order effects of the self-force are incorporated by solving the slow-time equation and letting $\tilde{t}$ vary. (Solving the next-higher-order slow-time equation determines similar effects, but also the backreaction that causes the parameters of the large black hole to change slowly.)

Using this method, Hinderer and Flanagan identified the effects of the various pieces of the self-force. To describe this we write the self-force as

$$f^\mu = \frac{m}{M}\left(f^\mu_{(1)\mathrm{rr}} + f^\mu_{(1)\mathrm{c}}\right) + \frac{m^2}{M^2}\left(f^\mu_{(2)\mathrm{rr}} + f^\mu_{(2)\mathrm{c}}\right) + \cdots, \tag{2.36}$$

where 'rr' denotes a radiation-reaction, or dissipative, piece of the force, and 'c' denotes a conservative piece. Hinderer and Flanagan's principal result is a formula for the orbital phase (which directly determines the phase of the emitted gravitational waves) in terms of these quantities:

$$\phi = \frac{M^2}{m}\left(\phi^{(0)}(\tilde{t}) + \frac{m}{M}\phi^{(1)}(\tilde{t}) + \cdots\right), \tag{2.37}$$

where $\phi^{(0)}$ depends on an averaged piece of $f^\mu_{(1)\mathrm{rr}}$, while $\phi^{(1)}$ depends on $f^\mu_{(1)\mathrm{c}}$, the oscillatory piece of $f^\mu_{(1)\mathrm{rr}}$, and the averaged piece of $f^\mu_{(2)\mathrm{rr}}$. From this result, we see that the radiative approximation yields the leading-order phase, but fails to determine the first subleading correction. We also see that the approximations (i)–(iii) mentioned above are consistent (so long as the parameters of the 'geodesic' source are allowed to vary slowly) at leading order in the two-timescale expansion, but diverge from one another beyond that order. Hence, we see that an adiabatic approximation is generically insufficient to extract parameters from a waveform, since doing so requires a description of the inspiral accurate up to small (i.e., smaller than order-1) errors. But we also see that an adiabatic approximation based on the radiative Green's function may be an excellent approximation for other purposes, such as detection.

To understand this result, consider the following naive analysis of a quasicircular EMRI — that is, an orbit that would be circular were it not for the action of the self-force, and which is slowly spiraling into the large central body. We write the orbital frequency as $\omega^{(0)}(E) + (m/M)\omega_1^{(1)}(E) + \cdots$, where $\omega^{(0)}(E)$ is the frequency as a function of energy on a circular geodesic, and $(m/M)\omega_1^{(1)}(E)$ is the correction to this due to the conservative part of the first-order self-force (part of the correction also arises due to oscillatory zeroth-order effects combining with oscillatory first-order effects, but for simplicity we ignore this contribution). Neglecting oscillatory effects, we write the energy in terms only of its slow-time dependence: $E = E^{(0)}(\tilde{t}) + (m/M)E^{(1)}(\tilde{t}) + \cdots$. The leading-order term $E^{(0)}$ is determined by the dissipative part of first-order self-force, while $E^{(1)}$ is determined by both the dissipative part of the second-order force and a combination of conservative and dissipative parts of the first-order force. Substituting this into the frequency, we arrive at

$$\omega = \omega^{(0)}(E^{(0)}) + \frac{m}{M}\left[\omega_1^{(1)}(E^{(0)}) + \omega_2^{(1)}(E^{(0)}, E^{(1)})\right] + \cdots, \tag{2.38}$$

where $\omega_2^{(1)} = E^{(1)}\partial\omega^{(0)}/\partial E$, in which the partial derivative is evaluated at $E = E^{(0)}$. Integrating this over a radiation-reaction time, we arrive at the orbital phase of Eq. (2.37). (In a complete description, $E(t)$ will have oscillatory pieces, which are functions of t rather than $\tilde{t}$, and one must know these in order to correctly determine $\phi^{(1)}$.) Such a result remains valid even for generic orbits, where, for example, orbital precession due to the conservative force contributes to the analogue of $\omega_1^{(1)}$.

2.6 Physical consequences of the self-force

To be of relevance to gravitational-wave astronomy, the paramount goal of the self-force community remains the computation of waveforms that properly encode the long-term dynamical evolution of an extreme-mass-ratio binary. This requires a fully consistent orbital evolution fed to a wave-generation formalism, and

to this day the completion of this program remains as a future challenge. In the meantime, a somewhat less ambitious, though no less compelling, undertaking is that of probing the physical consequences of the self-force on the motion of point particles.

Scalar charge in cosmological spacetimes

The intriguing phenomenon of a scalar charge changing its rest mass because of an interaction with its self-field was studied by Burko, Harte, and Poisson [47] and Haas and Poisson [48] in the simple context of a particle at rest in an expanding universe. The scalar Green's function could be computed explicitly for a wide class of cosmological spacetimes, and the action of the field on the particle determined without approximations. It is found that for certain cosmological models, the mass decreases and then increases back to its original value. For other models, the mass is restored only to a fraction of its original value. For de Sitter spacetime, the particle radiates all of its rest mass into monopole scalar waves.

Physical consequences of the gravitational self-force

The earliest calculation of a gravitational self-force was performed by Barack and Lousto for the case of a point mass plunging radially into a Schwarzschild black hole [132]. The calculation, however, depended on a specific choice of gauge and did not identify unambiguous physical consequences of the self-force. To obtain such consequences, it is necessary to combine the self-force (computed in whatever gauge) with the metric perturbation (computed in the same gauge) in the calculation of a well-defined observable that could in principle be measured. For example, the conservative pieces of the self-force and metric perturbation can be combined to calculate the shifts in orbital frequencies that originate from the gravitational effects of the small body; an application of such a calculation would be to determine the shift (as measured by frequency) in the innermost stable circular orbit of an extreme-mass-ratio binary, or the shift in the rate of periastron advance for eccentric orbits. Such calculations, however, must exclude all dissipative aspects of the self-force, because these introduce an inherent ambiguity in the determination of orbital frequencies.

A calculation of this kind was recently achieved by Barack and Sago [60, 133], who computed the shift in the innermost stable circular orbit of a Schwarzschild black hole caused by the conservative piece of the gravitational self-force. The shift in orbital radius is gauge dependent (and was reported in the Lorenz gauge by Barack and Sago), but the shift in orbital frequency is measurable and therefore gauge invariant. Their main result — a genuine milestone in self-force computations — is that the fractional shift in frequency is equal to $0.4870m/M$; the frequency is driven upward by the gravitational self-force. Barack and Sago compare this shift to the ambiguity created by the dissipative piece of the self-force, which was previously investigated by Ori and Thorne [134] and Sundararajan [135]; they find that the conservative shift is very small compared with the dissipative ambiguity. In a follow-up analysis, Barack, Damour, and Sago [63] computed the conservative shift in the rate of periastron advance of slightly eccentric orbits in Schwarzschild spacetime.

Conservative shifts in the innermost stable circular orbit of a Schwarzschild black hole were first obtained in the context of the scalar self-force by Diaz-Rivera *et al.* [136]; in this case they obtain a fractional shift of $0.0291657q^2/(mM)$, and here also the frequency is driven upward.

Detweiler's redshift factor

In another effort to extract physical consequences from the gravitational self-force on a particle in circular motion in Schwarzschild spacetime, Detweiler discovered [61] that u^t, the time component of the velocity vector in Schwarzschild coordinates, is invariant with respect to a class of gauge transformations that preserve the helical symmetry of the perturbed spacetime. Detweiler further showed that $1/u^t$ is an observable: it is the redshift that a photon suffers when it propagates from the orbiting body to an observer situated at a large distance on the orbital axis. This gauge-invariant quantity can be calculated together with the orbital frequency Ω, which is a second gauge-invariant quantity that can be constructed for circular orbits in Schwarzschild spacetime. Both u^t and Ω acquire corrections of fractional order m/M from the self-force and the metric perturbation. While the functions $u^t(r)$ and $\Omega(r)$ are still gauge dependent, because of the dependence on the radial coordinate r, elimination of r from these relations permits the construction of

$u^t(\Omega)$, which is gauge invariant. A plot of u^t as a function of Ω therefore contains physically unambiguous information about the gravitational self-force.

The computation of the gauge-invariant relation $u^t(\Omega)$ opened the door to a detailed comparison between the predictions of the self-force formalism to those of post-Newtonian theory. This was first pursued by Detweiler [61], who compared $u^t(\Omega)$ as determined accurately through second post-Newtonian order, to self-force results obtained numerically; he reported full consistency at the expected level of accuracy. This comparison was pushed to the third post-Newtonian order [63–66]. Agreement is remarkable, and it conveys a rather deep point about the methods of calculation. The computation of $u^t(\Omega)$, in the context of both the self-force and post-Newtonian theory, requires regularization of the metric perturbation created by the point mass. In the self-force calculation, removal of the singular field is achieved with the Detweiler-Whiting prescription, while in post-Newtonian theory it is performed with a very different prescription based on dimensional regularization. Each prescription could have returned a different regularized field, and therefore a different expression for $u^t(\Omega)$. This, remarkably, does not happen; the singular fields are "physically the same" in the self-force and post-Newtonian calculations.

A generalization of Detweiler's redshift invariant to eccentric orbits was recently proposed and computed by Barack and Sago [137], who report consistency with corresponding post-Newtonian results in the weak-field regime. They also computed the influence of the conservative gravitational self-force on the periastron advance of slightly eccentric orbits, and compared their results with full numerical relativity simulations for modest mass-ratio binaries. Thus, in spite of the unavailability of self-consistent waveforms, it is becoming clear that self-force calculations are already proving to be of value: they inform post-Newtonian calculations and serve as benchmarks for numerical relativity.

Part I
General theory of bitensors

3 Synge's world function

3.1 Definition

In this and the following sections we will construct a number of *bitensors*, tensorial functions of two points in spacetime. The first is x', which we call the "base point", and to which we assign indices α', β', etc. The second is x, which we call the "field point", and to which we assign indices α, β, etc. We assume that x belongs to $\mathcal{N}(x')$, the *normal convex neighbourhood* of x'; this is the set of points that are linked to x' by a *unique* geodesic. The geodesic segment β that links x to x' is described by relations $z^\mu(\lambda)$ in which λ is an affine parameter that ranges from λ_0 to λ_1; we have $z(\lambda_0) := x'$ and $z(\lambda_1) := x$. To an arbitrary point z on the geodesic we assign indices μ, ν, etc. The vector $t^\mu = dz^\mu/d\lambda$ is tangent to the geodesic, and it obeys the geodesic equation $Dt^\mu/d\lambda = 0$. The situation is illustrated in Fig. 5.

Synge's world function is a scalar function of the base point x' and the field point x. It is defined by

$$\sigma(x, x') = \frac{1}{2}(\lambda_1 - \lambda_0) \int_{\lambda_0}^{\lambda_1} g_{\mu\nu}(z) t^\mu t^\nu \, d\lambda, \tag{3.1}$$

and the integral is evaluated on the geodesic β that links x to x'. You may notice that σ is invariant under a constant rescaling of the affine parameter, $\lambda \to \bar{\lambda} = a\lambda + b$, where a and b are constants.

By virtue of the geodesic equation, the quantity $\varepsilon := g_{\mu\nu} t^\mu t^\nu$ is constant on the geodesic. The world function is therefore numerically equal to $\frac{1}{2}\varepsilon(\lambda_1 - \lambda_0)^2$. If the geodesic is timelike, then λ can be set equal to the proper time τ, which implies that $\varepsilon = -1$ and $\sigma = -\frac{1}{2}(\Delta\tau)^2$. If the geodesic is spacelike, then λ can be set equal to the proper distance s, which implies that $\varepsilon = 1$ and $\sigma = \frac{1}{2}(\Delta s)^2$. If the geodesic is null, then $\sigma = 0$. Quite generally, therefore, the world function is half the squared geodesic distance between the points x' and x.

In flat spacetime, the geodesic linking x to x' is a straight line, and $\sigma = \frac{1}{2}\eta_{\alpha\beta}(x - x')^\alpha (x - x')^\beta$ in Lorentzian coordinates.

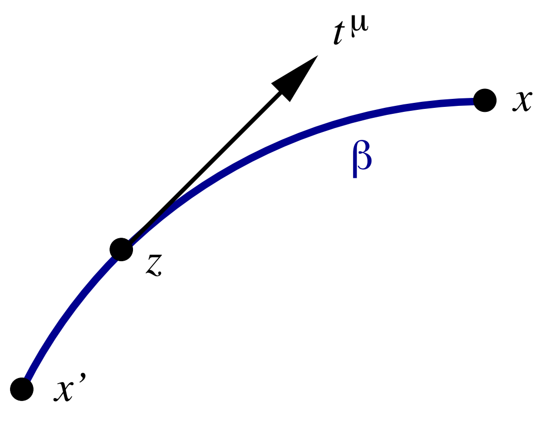

Figure 5: The base point x', the field point x, and the geodesic segment β that links them. The geodesic is described by parametric relations $z^\mu(\lambda)$ and $t^\mu = dz^\mu/d\lambda$ is its tangent vector.

3.2 Differentiation of the world function

The world function $\sigma(x, x')$ can be differentiated with respect to either argument. We let $\sigma_\alpha = \partial\sigma/\partial x^\alpha$ be its partial derivative with respect to x, and $\sigma_{\alpha'} = \partial\sigma/\partial x^{\alpha'}$ its partial derivative with respect to x'. It is clear that σ_α behaves as a dual vector with respect to tensorial operations carried out at x, but as a scalar with respect to operations carried out x'. Similarly, $\sigma_{\alpha'}$ is a scalar at x but a dual vector at x'.

We let $\sigma_{\alpha\beta} := \nabla_\beta\sigma_\alpha$ be the covariant derivative of σ_α with respect to x; this is a rank-2 tensor at x and a scalar at x'. Because σ is a scalar at x, we have that this tensor is symmetric: $\sigma_{\beta\alpha} = \sigma_{\alpha\beta}$. Similarly, we let $\sigma_{\alpha\beta'} := \partial_{\beta'}\sigma_\alpha = \partial^2\sigma/\partial x^{\beta'}\partial x^\alpha$ be the partial derivative of σ_α with respect to x'; this is a dual vector both at x and x'. We can also define $\sigma_{\alpha'\beta} := \partial_\beta\sigma_{\alpha'} = \partial^2\sigma/\partial x^\beta\partial x^{\alpha'}$ to be the partial derivative of $\sigma_{\alpha'}$ with respect to x. Because partial derivatives commute, these bitensors are equal: $\sigma_{\beta'\alpha} = \sigma_{\alpha\beta'}$. Finally, we let $\sigma_{\alpha'\beta'} := \nabla_{\beta'}\sigma_{\alpha'}$ be the covariant derivative of $\sigma_{\alpha'}$ with respect to x'; this is a symmetric rank-2 tensor at x' and a scalar at x.

The notation is easily extended to any number of derivatives. For example, we let $\sigma_{\alpha\beta\gamma\delta'} := \nabla_{\delta'}\nabla_\gamma\nabla_\beta\nabla_\alpha\sigma$, which is a rank-3 tensor at x and a dual vector at x'. This bitensor is symmetric in the pair of indices α and β, but not in the pairs α and γ, nor β and γ. Because $\nabla_{\delta'}$ is here an ordinary partial derivative with respect to x', the bitensor is symmetric in any pair of indices involving δ'. The ordering of the primed index relative to the unprimed indices is therefore irrelevant: the same bitensor can be written as $\sigma_{\delta'\alpha\beta\gamma}$ or $\sigma_{\alpha\delta'\beta\gamma}$ or $\sigma_{\alpha\beta\delta'\gamma}$, making sure that the ordering of the unprimed indices is not altered.

More generally, we can show that derivatives of any bitensor $\Omega_{\cdots}(x, x')$ satisfy the property

$$\Omega_{\cdots;\beta\alpha'\cdots} = \Omega_{\cdots;\alpha'\beta\cdots}, \tag{3.2}$$

in which "$\cdots$" stands for any combination of primed and unprimed indices. We start by establishing the symmetry of $\Omega_{\cdots;\alpha\beta'}$ with respect to the pair α and β'. This is most easily done by adopting Fermi normal coordinates (see Sec. 9) adapted to the geodesic β and setting the connection to zero both at x and x'. In these coordinates, the bitensor $\Omega_{\cdots;\alpha}$ is the partial derivative of $\Omega_{\cdots}$ with respect to x^α, and $\Omega_{\cdots;\alpha\beta'}$ is obtained by taking an additional partial derivative with respect to $x^{\beta'}$. These two operations commute, and $\Omega_{\cdots;\beta'\alpha} = \Omega_{\cdots;\alpha\beta'}$ follows as a bitensorial identity. Equation (3.2) then follows by further differentiation with respect to either x or x'.

The message of Eq. (3.2), when applied to derivatives of the world function, is that while the ordering of the primed and unprimed indices relative to themselves is important, their ordering with respect to each other is arbitrary. For example, $\sigma_{\alpha'\beta'\gamma\delta'\epsilon} = \sigma_{\alpha'\beta'\delta'\gamma\epsilon} = \sigma_{\gamma\epsilon\alpha'\beta'\delta'}$.

3.3 Evaluation of first derivatives

We can compute σ_α by examining how σ varies when the field point x moves. We let the new field point be $x + \delta x$, and $\delta\sigma := \sigma(x + \delta x, x') - \sigma(x, x')$ is the corresponding variation of the world function. We let $\beta + \delta\beta$ be the unique geodesic segment that links $x + \delta x$ to x'; it is described by relations $z^\mu(\lambda) + \delta z^\mu(\lambda)$, in which the affine parameter is scaled in such a way that it runs from λ_0 to λ_1 also on the new geodesic. We note that $\delta z(\lambda_0) = \delta x' = 0$ and $\delta z(\lambda_1) = \delta x$.

Working to first order in the variations, Eq. (3.1) implies

$$\delta\sigma = \Delta\lambda \int_{\lambda_0}^{\lambda_1} \Bigl(g_{\mu\nu}\dot{z}^\mu\,\delta\dot{z}^\nu + \frac{1}{2}\,g_{\mu\nu,\lambda}\dot{z}^\mu\dot{z}^\nu\,\delta z^\lambda \Bigr)\,d\lambda,$$

where $\Delta\lambda = \lambda_1 - \lambda_0$, an overdot indicates differentiation with respect to λ, and the metric and its derivatives are evaluated on β. Integrating the first term by parts gives

$$\delta\sigma = \Delta\lambda\Bigl[g_{\mu\nu}\dot{z}^\mu\,\delta z^\nu\Bigr]_{\lambda_0}^{\lambda_1} - \Delta\lambda\int_{\lambda_0}^{\lambda_1}\Bigl(g_{\mu\nu}\ddot{z}^\nu + \Gamma_{\mu\nu\lambda}\dot{z}^\nu\dot{z}^\lambda\Bigr)\,\delta z^\mu\,d\lambda.$$

The integral vanishes because $z^\mu(\lambda)$ satisfies the geodesic equation. The boundary term at λ_0 is zero because the variation δz^μ vanishes there. We are left with $\delta\sigma = \Delta\lambda g_{\alpha\beta}t^\alpha\delta x^\beta$, or

$$\sigma_\alpha(x, x') = (\lambda_1 - \lambda_0)\,g_{\alpha\beta}t^\beta, \tag{3.3}$$

in which the metric and the tangent vector are both evaluated at x. Apart from a factor $\Delta\lambda$, we see that $\sigma^\alpha(x,x')$ is equal to the geodesic's tangent vector at x. If in Eq. (3.3) we replace x by a generic point $z(\lambda)$ on β, and if we correspondingly replace λ_1 by λ, we obtain $\sigma^\mu(z,x') = (\lambda-\lambda_0)t^\mu$; we therefore see that $\sigma^\mu(z,x')$ is a rescaled tangent vector on the geodesic.

A virtually identical calculation reveals how σ varies under a change of base point x'. Here the variation of the geodesic is such that $\delta z(\lambda_0) = \delta x'$ and $\delta z(\lambda_1) = \delta x = 0$, and we obtain $\delta\sigma = -\Delta\lambda g_{\alpha'\beta'}t^{\alpha'}\delta x^{\beta'}$. This shows that

$$\sigma_{\alpha'}(x,x') = -(\lambda_1-\lambda_0)\, g_{\alpha'\beta'}t^{\beta'}, \tag{3.4}$$

in which the metric and the tangent vector are both evaluated at x'. Apart from a factor $\Delta\lambda$, we see that $\sigma^{\alpha'}(x,x')$ is minus the geodesic's tangent vector at x'.

It is interesting to compute the norm of σ_α. According to Eq. (3.3) we have $g_{\alpha\beta}\sigma^\alpha\sigma^\beta = (\Delta\lambda)^2 g_{\alpha\beta}t^\alpha t^\beta = (\Delta\lambda)^2\varepsilon$. According to Eq. (3.1), this is equal to 2σ. We have obtained

$$g^{\alpha\beta}\sigma_\alpha\sigma_\beta = 2\sigma, \tag{3.5}$$

and similarly,

$$g^{\alpha'\beta'}\sigma_{\alpha'}\sigma_{\beta'} = 2\sigma. \tag{3.6}$$

These important relations will be the starting point of many computations to be described below.

We note that in flat spacetime, $\sigma_\alpha = \eta_{\alpha\beta}(x-x')^\beta$ and $\sigma_{\alpha'} = -\eta_{\alpha\beta}(x-x')^\beta$ in Lorentzian coordinates. From this it follows that $\sigma_{\alpha\beta} = \sigma_{\alpha'\beta'} = -\sigma_{\alpha\beta'} = -\sigma_{\alpha'\beta} = \eta_{\alpha\beta}$, and finally, $g^{\alpha\beta}\sigma_{\alpha\beta} = 4 = g^{\alpha'\beta'}\sigma_{\alpha'\beta'}$.

3.4 Congruence of geodesics emanating from x'

If the base point x' is kept fixed, σ can be considered to be an ordinary scalar function of x. According to Eq. (3.5), this function is a solution to the nonlinear differential equation $\frac{1}{2}g^{\alpha\beta}\sigma_\alpha\sigma_\beta = \sigma$. Suppose that we are presented with such a scalar field. What can we say about it?

An additional differentiation of the defining equation reveals that the vector $\sigma^\alpha := \sigma^{;\alpha}$ satisfies

$$\sigma^\alpha{}_{;\beta}\sigma^\beta = \sigma^\alpha, \tag{3.7}$$

which is the geodesic equation in a non-affine parameterization. The vector field is therefore tangent to a congruence of geodesics. The geodesics are timelike where $\sigma < 0$, they are spacelike where $\sigma > 0$, and they are null where $\sigma = 0$. Here, for concreteness, we shall consider only the timelike subset of the congruence.

The vector

$$u^\alpha = \frac{\sigma^\alpha}{|2\sigma|^{1/2}} \tag{3.8}$$

is a normalized tangent vector that satisfies the geodesic equation in affine-parameter form: $u^\alpha{}_{;\beta}u^\beta = 0$. The parameter λ is then proper time τ. If λ^* denotes the original parameterization of the geodesics, we have that $d\lambda^*/d\tau = |2\sigma|^{-1/2}$, and we see that the original parameterization is singular at $\sigma = 0$.

In the affine parameterization, the expansion of the congruence is calculated to be

$$\theta = \frac{\theta^*}{|2\sigma|^{1/2}}, \qquad \theta^* := \sigma^\alpha{}_{;\alpha} - 1, \tag{3.9}$$

where $\theta^* = (\delta V)^{-1}(d/d\lambda^*)(\delta V)$ is the expansion in the original parameterization (δV is the congruence's cross-sectional volume). While θ^* is well behaved in the limit $\sigma \to 0$ (we shall see below that $\theta^* \to 3$), we have that $\theta \to \infty$. This means that the point x' at which $\sigma = 0$ is a caustic of the congruence: all geodesics emanate from this point.

These considerations, which all follow from a postulated relation $\frac{1}{2}g^{\alpha\beta}\sigma_\alpha\sigma_\beta = \sigma$, are clearly compatible with our preceding explicit construction of the world function.

4 Coincidence limits

It is useful to determine the limiting behaviour of the bitensors $\sigma_{\dots}$ as x approaches x'. We introduce the notation

$$\big[\Omega_{\dots}\big] = \lim_{x\to x'} \Omega_{\dots}(x,x') = \text{a tensor at } x'$$

to designate the limit of any bitensor $\Omega_{\dots}(x,x')$ as x approaches x'; this is called the *coincidence limit* of the bitensor. We assume that the coincidence limit is a unique tensorial function of the base point x', independent of the direction in which the limit is taken. In other words, if the limit is computed by letting $\lambda \to \lambda_0$ after evaluating $\Omega_{\dots}(z,x')$ as a function of λ on a specified geodesic β, it is assumed that the answer does not depend on the choice of geodesic.

4.1 Computation of coincidence limits

From Eqs. (3.1), (3.3), and (3.4) we already have

$$\big[\sigma\big] = 0, \qquad \big[\sigma_\alpha\big] = \big[\sigma_{\alpha'}\big] = 0. \tag{4.1}$$

Additional results are obtained by repeated differentiation of the relations (3.5) and (3.6). For example, Eq. (3.5) implies $\sigma_\gamma = g^{\alpha\beta}\sigma_\alpha\sigma_{\beta\gamma} = \sigma^\beta\sigma_{\beta\gamma}$, or $(g_{\beta\gamma} - \sigma_{\beta\gamma})t^\beta = 0$ after using Eq. (3.3). From the assumption stated in the preceding paragraph, $\sigma_{\beta\gamma}$ becomes independent of t^β in the limit $x \to x'$, and we arrive at $[\sigma_{\alpha\beta}] = g_{\alpha'\beta'}$. By very similar calculations we obtain all other coincidence limits for the second derivatives of the world function. The results are

$$\big[\sigma_{\alpha\beta}\big] = \big[\sigma_{\alpha'\beta'}\big] = g_{\alpha'\beta'}, \qquad \big[\sigma_{\alpha\beta'}\big] = \big[\sigma_{\alpha'\beta}\big] = -g_{\alpha'\beta'}. \tag{4.2}$$

From these relations we infer that $[\sigma^\alpha_{\ \alpha}] = 4$, so that $[\theta^*] = 3$, where θ^* was defined in Eq. (3.9).

To generate coincidence limits of bitensors involving primed indices, it is efficient to invoke Synge's rule,

$$\big[\sigma_{\dots\alpha'}\big] = \big[\sigma_{\dots}\big]_{;\alpha'} - \big[\sigma_{\dots\alpha}\big], \tag{4.3}$$

in which "$\cdots$" designates any combination of primed and unprimed indices; this rule will be established below. For example, according to Synge's rule we have $[\sigma_{\alpha\beta'}] = [\sigma_\alpha]_{;\beta'} - [\sigma_{\alpha\beta}]$, and since the coincidence limit of σ_α is zero, this gives us $[\sigma_{\alpha\beta'}] = -[\sigma_{\alpha\beta}] = -g_{\alpha'\beta'}$, as was stated in Eq. (4.2). Similarly, $[\sigma_{\alpha'\beta'}] = [\sigma_{\alpha'}]_{;\beta'} - [\sigma_{\alpha'\beta}] = -[\sigma_{\beta\alpha'}] = g_{\alpha'\beta'}$. The results of Eq. (4.2) can thus all be generated from the known result for $[\sigma_{\alpha\beta}]$.

The coincidence limits of Eq. (4.2) were derived from the relation $\sigma_\alpha = \sigma^\delta_{\ \alpha}\sigma_\delta$. We now differentiate this twice more and obtain $\sigma_{\alpha\beta\gamma} = \sigma^\delta_{\ \alpha\beta\gamma}\sigma_\delta + \sigma^\delta_{\ \alpha\beta}\sigma_{\delta\gamma} + \sigma^\delta_{\ \alpha\gamma}\sigma_{\delta\beta} + \sigma^\delta_{\ \alpha}\sigma_{\delta\beta\gamma}$. At coincidence we have

$$\big[\sigma_{\alpha\beta\gamma}\big] = \big[\sigma^\delta_{\ \alpha\beta}\big]g_{\delta'\gamma'} + \big[\sigma^\delta_{\ \alpha\gamma}\big]g_{\delta'\beta'} + \delta^{\delta'}_{\ \alpha'}\big[\sigma_{\delta\beta\gamma}\big],$$

or $[\sigma_{\gamma\alpha\beta}] + [\sigma_{\beta\alpha\gamma}] = 0$ if we recognize that the operations of raising or lowering indices and taking the limit $x \to x'$ commute. Noting the symmetries of $\sigma_{\alpha\beta}$, this gives us $[\sigma_{\alpha\gamma\beta}] + [\sigma_{\alpha\beta\gamma}] = 0$, or $2[\sigma_{\alpha\beta\gamma}] - [R^\delta_{\ \alpha\beta\gamma}\sigma_\delta] = 0$, or $2[\sigma_{\alpha\beta\gamma}] = R^{\delta'}_{\ \alpha'\beta'\gamma'}[\sigma_{\delta'}]$. Since the last factor is zero, we arrive at

$$\big[\sigma_{\alpha\beta\gamma}\big] = \big[\sigma_{\alpha\beta\gamma'}\big] = \big[\sigma_{\alpha\beta'\gamma'}\big] = \big[\sigma_{\alpha'\beta'\gamma'}\big] = 0. \tag{4.4}$$

The last three results were derived from $[\sigma_{\alpha\beta\gamma}] = 0$ by employing Synge's rule.

We now differentiate the relation $\sigma_\alpha = \sigma^\delta_{\ \alpha}\sigma_\delta$ three times and obtain

$$\sigma_{\alpha\beta\gamma\delta} = \sigma^\epsilon_{\ \alpha\beta\gamma\delta}\sigma_\epsilon + \sigma^\epsilon_{\ \alpha\beta\gamma}\sigma_{\epsilon\delta} + \sigma^\epsilon_{\ \alpha\beta\delta}\sigma_{\epsilon\gamma} + \sigma^\epsilon_{\ \alpha\gamma\delta}\sigma_{\epsilon\beta} + \sigma^\epsilon_{\ \alpha\beta}\sigma_{\epsilon\gamma\delta} + \sigma^\epsilon_{\ \alpha\gamma}\sigma_{\epsilon\beta\delta} + \sigma^\epsilon_{\ \alpha\delta}\sigma_{\epsilon\beta\gamma} + \sigma^\epsilon_{\ \alpha}\sigma_{\epsilon\beta\gamma\delta}.$$

At coincidence this reduces to $[\sigma_{\alpha\beta\gamma\delta}] + [\sigma_{\alpha\delta\beta\gamma}] + [\sigma_{\alpha\gamma\beta\delta}] = 0$. To simplify the third term we differentiate Ricci's identity $\sigma_{\alpha\gamma\beta} = \sigma_{\alpha\beta\gamma} - R^\epsilon_{\ \alpha\beta\gamma}\sigma_\epsilon$ with respect to x^δ and then take the coincidence limit. This gives us $[\sigma_{\alpha\gamma\beta\delta}] = [\sigma_{\alpha\beta\gamma\delta}] + R_{\alpha'\delta'\beta'\gamma'}$. The same manipulations on the second term give $[\sigma_{\alpha\delta\beta\gamma}] = [\sigma_{\alpha\beta\delta\gamma}] + R_{\alpha'\gamma'\beta'\delta'}$. Using the identity $\sigma_{\alpha\beta\delta\gamma} = \sigma_{\alpha\beta\gamma\delta} - R^\epsilon_{\ \alpha\gamma\delta}\sigma_{\epsilon\beta} - R^\epsilon_{\ \beta\gamma\delta}\sigma_{\alpha\epsilon}$ and the symmetries of the Riemann tensor, it is then easy to show that $[\sigma_{\alpha\beta\delta\gamma}] = [\sigma_{\alpha\beta\gamma\delta}]$. Gathering the results, we obtain $3[\sigma_{\alpha\beta\gamma\delta}] + R_{\alpha'\gamma'\beta'\delta'} + R_{\alpha'\delta'\beta'\gamma'} = 0$,

and Synge's rule allows us to generalize this to any combination of primed and unprimed indices. Our final results are

$$\begin{aligned}
\big[\sigma_{\alpha\beta\gamma\delta}\big] &= -\frac{1}{3}\big(R_{\alpha'\gamma'\beta'\delta'} + R_{\alpha'\delta'\beta'\gamma'}\big), \qquad \big[\sigma_{\alpha\beta\gamma\delta'}\big] = \frac{1}{3}\big(R_{\alpha'\gamma'\beta'\delta'} + R_{\alpha'\delta'\beta'\gamma'}\big),\\
\big[\sigma_{\alpha\beta\gamma'\delta'}\big] &= -\frac{1}{3}\big(R_{\alpha'\gamma'\beta'\delta'} + R_{\alpha'\delta'\beta'\gamma'}\big), \qquad \big[\sigma_{\alpha\beta'\gamma'\delta'}\big] = -\frac{1}{3}\big(R_{\alpha'\beta'\gamma'\delta'} + R_{\alpha'\gamma'\beta'\delta'}\big),\\
\big[\sigma_{\alpha'\beta'\gamma'\delta'}\big] &= -\frac{1}{3}\big(R_{\alpha'\gamma'\beta'\delta'} + R_{\alpha'\delta'\beta'\gamma'}\big).
\end{aligned} \tag{4.5}$$

4.2 Derivation of Synge's rule

We begin with *any* bitensor $\Omega_{AB'}(x,x')$ in which $A = \alpha\cdots\beta$ is a multi-index that represents any number of unprimed indices, and $B' = \gamma'\cdots\delta'$ a multi-index that represents any number of primed indices. (It does not matter whether the primed and unprimed indices are segregated or mixed.) On the geodesic β that links x to x' we introduce an ordinary tensor $P^M(z)$ where M is a multi-index that contains the same number of indices as A. This tensor is arbitrary, but we assume that it is parallel transported on β; this means that it satisfies $P^A{}_{;\alpha}t^\alpha = 0$ at x. Similarly, we introduce an ordinary tensor $Q^N(z)$ in which N contains the same number of indices as B'. This tensor is arbitrary, but we assume that it is parallel transported on β; at x' it satisfies $Q^{B'}{}_{;\alpha'}t^{\alpha'} = 0$. With Ω, P, and Q we form a biscalar $H(x,x')$ defined by

$$H(x,x') = \Omega_{AB'}(x,x')P^A(x)Q^{B'}(x').$$

Having specified the geodesic that links x to x', we can consider H to be a function of λ_0 and λ_1. If λ_1 is not much larger than λ_0 (so that x is not far from x'), we can express $H(\lambda_1,\lambda_0)$ as

$$H(\lambda_1,\lambda_0) = H(\lambda_0,\lambda_0) + (\lambda_1-\lambda_0)\,\frac{\partial H}{\partial\lambda_1}\bigg|_{\lambda_1=\lambda_0} + \cdots.$$

Alternatively,

$$H(\lambda_1,\lambda_0) = H(\lambda_1,\lambda_1) - (\lambda_1-\lambda_0)\,\frac{\partial H}{\partial\lambda_0}\bigg|_{\lambda_0=\lambda_1} + \cdots,$$

and these two expressions give

$$\frac{d}{d\lambda_0}H(\lambda_0,\lambda_0) = \frac{\partial H}{\partial\lambda_0}\bigg|_{\lambda_0=\lambda_1} + \frac{\partial H}{\partial\lambda_1}\bigg|_{\lambda_1=\lambda_0},$$

because the left-hand side is the limit of $[H(\lambda_1,\lambda_1) - H(\lambda_0,\lambda_0)]/(\lambda_1-\lambda_0)$ when $\lambda_1\to\lambda_0$. The partial derivative of H with respect to λ_0 is equal to $\Omega_{AB';\alpha'}t^{\alpha'}P^AQ^{B'}$, and in the limit this becomes $[\Omega_{AB';\alpha'}]t^{\alpha'}P^{A'}Q^{B'}$. Similarly, the partial derivative of H with respect to λ_1 is $\Omega_{AB';\alpha}t^\alpha P^AQ^{B'}$, and in the limit $\lambda_1\to\lambda_0$ this becomes $[\Omega_{AB';\alpha}]t^{\alpha'}P^{A'}Q^{B'}$. Finally, $H(\lambda_0,\lambda_0) = [\Omega_{AB'}]P^{A'}Q^{B'}$, and its derivative with respect to λ_0 is $[\Omega_{AB'}]_{;\alpha'}t^{\alpha'}P^{A'}Q^{B'}$. Gathering the results we find that

$$\Big\{\big[\Omega_{AB'}\big]_{;\alpha'} - \big[\Omega_{AB';\alpha'}\big] - \big[\Omega_{AB';\alpha}\big]\Big\}t^{\alpha'}P^{A'}Q^{B'} = 0,$$

and the final statement of Synge's rule,

$$\big[\Omega_{AB'}\big]_{;\alpha'} = \big[\Omega_{AB';\alpha'}\big] + \big[\Omega_{AB';\alpha}\big], \tag{4.6}$$

follows from the fact that the tensors P^M and Q^N, and the direction of the selected geodesic β, are all arbitrary. Equation (4.6) reduces to Eq. (4.3) when $\sigma_{\dots}$ is substituted in place of $\Omega_{AB'}$.

5 Parallel propagator

5.1 Tetrad on β

On the geodesic segment β that links x to x' we introduce an orthonormal basis $e^\mu_{\mathsf{a}}(z)$ that is parallel transported on the geodesic. The frame indices a, b, $\dots$, run from 0 to 3 and the basis vectors satisfy

$$g_{\mu\nu}\, e^\mu_{\mathsf{a}} e^\nu_{\mathsf{b}} = \eta_{\mathsf{ab}}, \qquad \frac{De^\mu_{\mathsf{a}}}{d\lambda} = 0, \tag{5.1}$$

where $\eta_{\mathsf{ab}} = \mathrm{diag}(-1,1,1,1)$ is the Minkowski metric (which we shall use to raise and lower frame indices). We have the completeness relations

$$g^{\mu\nu} = \eta^{\mathsf{ab}}\, e^\mu_{\mathsf{a}} e^\nu_{\mathsf{b}}, \tag{5.2}$$

and we define a dual tetrad $e^{\mathsf{a}}_\mu(z)$ by

$$e^{\mathsf{a}}_\mu := \eta^{\mathsf{ab}} g_{\mu\nu}\, e^\nu_{\mathsf{b}}; \tag{5.3}$$

this is also parallel transported on β. In terms of the dual tetrad the completeness relations take the form

$$g_{\mu\nu} = \eta_{\mathsf{ab}}\, e^{\mathsf{a}}_\mu e^{\mathsf{b}}_\nu, \tag{5.4}$$

and it is easy to show that the tetrad and its dual satisfy $e^{\mathsf{a}}_\mu e^\mu_{\mathsf{b}} = \delta^{\mathsf{a}}{}_{\mathsf{b}}$ and $e^{\mathsf{a}}_\nu e^\mu_{\mathsf{a}} = \delta^\mu{}_\nu$. Equations (5.1)–(5.4) hold everywhere on β. In particular, with an appropriate change of notation they hold at x' and x; for example, $g_{\alpha\beta} = \eta_{\mathsf{ab}}\, e^{\mathsf{a}}_\alpha e^{\mathsf{b}}_\beta$ is the metric at x.

(You will have noticed that we use sans-serif symbols for the frame indices. This is to distinguish them from another set of frame indices that will appear below. The frame indices introduced here run from 0 to 3; those to be introduced later will run from 1 to 3.)

5.2 Definition and properties of the parallel propagator

Any vector field $A^\mu(z)$ on β can be decomposed in the basis e^μ_{a}: $A^\mu = A^{\mathsf{a}}\, e^\mu_{\mathsf{a}}$, and the vector's frame components are given by $A^{\mathsf{a}} = A^\mu\, e^{\mathsf{a}}_\mu$. If A^μ is parallel transported on the geodesic, then the coefficients A^{a} are constants. The vector at x can then be expressed as $A^\alpha = (A^{\alpha'} e^{\mathsf{a}}_{\alpha'}) e^\alpha_{\mathsf{a}}$, or

$$A^\alpha(x) = g^\alpha{}_{\alpha'}(x,x')\, A^{\alpha'}(x'), \qquad g^\alpha{}_{\alpha'}(x,x') := e^\alpha_{\mathsf{a}}(x)\, e^{\mathsf{a}}_{\alpha'}(x'). \tag{5.5}$$

The object $g^\alpha{}_{\alpha'} = e^\alpha_{\mathsf{a}} e^{\mathsf{a}}_{\alpha'}$ is the *parallel propagator*: it takes a vector at x' and parallel-transports it to x along the unique geodesic that links these points.

Similarly, we find that

$$A^{\alpha'}(x') = g^{\alpha'}{}_\alpha(x',x)\, A^\alpha(x), \qquad g^{\alpha'}{}_\alpha(x',x) := e^{\alpha'}_{\mathsf{a}}(x')\, e^{\mathsf{a}}_\alpha(x), \tag{5.6}$$

and we see that $g^{\alpha'}{}_\alpha = e^{\alpha'}_{\mathsf{a}} e^{\mathsf{a}}_\alpha$ performs the inverse operation: it takes a vector at x and parallel-transports it back to x'. Clearly,

$$g^\alpha{}_{\alpha'} g^{\alpha'}{}_\beta = \delta^\alpha{}_\beta, \qquad g^{\alpha'}{}_\alpha g^\alpha{}_{\beta'} = \delta^{\alpha'}{}_{\beta'}, \tag{5.7}$$

and these relations formally express the fact that $g^{\alpha'}{}_\alpha$ is the inverse of $g^\alpha{}_{\alpha'}$.

The relation $g^\alpha{}_{\alpha'} = e^\alpha_{\mathsf{a}} e^{\mathsf{a}}_{\alpha'}$ can also be expressed as $g_\alpha{}^{\alpha'} = e^{\mathsf{a}}_\alpha e^{\alpha'}_{\mathsf{a}}$, and this reveals that

$$g_\alpha{}^{\alpha'}(x,x') = g^{\alpha'}{}_\alpha(x',x), \qquad g_{\alpha'}{}^\alpha(x',x) = g^\alpha{}_{\alpha'}(x,x'). \tag{5.8}$$

The ordering of the indices, and the ordering of the arguments, are arbitrary.

The action of the parallel propagator on tensors of arbitrary rank is easy to figure out. For example, suppose that the dual vector $p_\mu = p_a\, e^a_\mu$ is parallel transported on β. Then the frame components $p_{\mathsf{a}} = p_\mu\, e^\mu_{\mathsf{a}}$ are constants, and the dual vector at x can be expressed as $p_\alpha = (p_{\alpha'} e^{\alpha'}_{\mathsf{a}}) e^\alpha_{\mathsf{a}}$, or

$$p_\alpha(x) = g^{\alpha'}{}_\alpha(x',x)\, p_{\alpha'}(x'). \tag{5.9}$$

It is therefore the inverse propagator $g^{\alpha'}_{\ \alpha}$ that takes a dual vector at x' and parallel-transports it to x. As another example, it is easy to show that a tensor $A^{\alpha\beta}$ at x obtained by parallel transport from x' must be given by

$$A^{\alpha\beta}(x) = g^{\alpha}_{\ \alpha'}(x,x') g^{\beta}_{\ \beta'}(x,x')\, A^{\alpha'\beta'}(x'). \tag{5.10}$$

Here we need two occurrences of the parallel propagator, one for each tensorial index. Because the metric tensor is covariantly constant, it is automatically parallel transported on β, and a special case of Eq. (5.10) is $g_{\alpha\beta} = g^{\alpha'}_{\ \alpha} g^{\beta'}_{\ \beta}\, g_{\alpha'\beta'}$.

Because the basis vectors are parallel transported on β, they satisfy $e^{\alpha}_{\mathsf{a};\beta}\sigma^{\beta} = 0$ at x and $e^{\alpha'}_{\mathsf{a};\beta'}\sigma^{\beta'} = 0$ at x'. This immediately implies that the parallel propagators must satisfy

$$g^{\alpha}_{\ \alpha';\beta}\sigma^{\beta} = g^{\alpha}_{\ \alpha';\beta'}\sigma^{\beta'} = 0, \qquad g^{\alpha'}_{\ \alpha;\beta}\sigma^{\beta} = g^{\alpha'}_{\ \alpha;\beta'}\sigma^{\beta'} = 0. \tag{5.11}$$

Another useful property of the parallel propagator follows from the fact that if $t^{\mu} = dz^{\mu}/d\lambda$ is tangent to the geodesic connecting x to x', then $t^{\alpha} = g^{\alpha}_{\ \alpha'} t^{\alpha'}$. Using Eqs. (3.3) and (3.4), this observation gives us the relations

$$\sigma_{\alpha} = -g^{\alpha'}_{\ \alpha}\sigma_{\alpha'}, \qquad \sigma_{\alpha'} = -g^{\alpha}_{\ \alpha'}\sigma_{\alpha}. \tag{5.12}$$

5.3 Coincidence limits

Equation (5.5) and the completeness relations of Eqs. (5.2) or (5.4) imply that

$$\big[g^{\alpha}_{\ \beta'}\big] = \delta^{\alpha'}_{\ \beta'}. \tag{5.13}$$

Other coincidence limits are obtained by differentiation of Eqs. (5.11). For example, the relation $g^{\alpha}_{\ \beta';\gamma}\sigma^{\gamma} = 0$ implies $g^{\alpha}_{\ \beta';\gamma\delta}\sigma^{\gamma} + g^{\alpha}_{\ \beta';\gamma}\sigma^{\gamma}_{\ \delta} = 0$, and at coincidence we have

$$\big[g^{\alpha}_{\ \beta';\gamma}\big] = \big[g^{\alpha}_{\ \beta';\gamma'}\big] = 0; \tag{5.14}$$

the second result was obtained by applying Synge's rule on the first result. Further differentiation gives

$$g^{\alpha}_{\ \beta';\gamma\delta\epsilon}\sigma^{\gamma} + g^{\alpha}_{\ \beta';\gamma\delta}\sigma^{\gamma}_{\ \epsilon} + g^{\alpha}_{\ \beta';\gamma\epsilon}\sigma^{\gamma}_{\ \delta} + g^{\alpha}_{\ \beta';\gamma}\sigma^{\gamma}_{\ \delta\epsilon} = 0,$$

and at coincidence we have $[g^{\alpha}_{\ \beta';\gamma\delta}] + [g^{\alpha}_{\ \beta';\delta\gamma}] = 0$, or $2[g^{\alpha}_{\ \beta';\gamma\delta}] + R^{\alpha'}_{\ \beta'\gamma'\delta'} = 0$. The coincidence limit for $g^{\alpha}_{\ \beta';\gamma\delta'} = g^{\alpha}_{\ \beta';\delta'\gamma}$ can then be obtained from Synge's rule, and an additional application of the rule gives $[g^{\alpha}_{\ \beta';\gamma'\delta'}]$. Our results are

$$\begin{aligned}
\big[g^{\alpha}_{\ \beta';\gamma\delta}\big] &= -\frac{1}{2}\,R^{\alpha'}_{\ \beta'\gamma'\delta'}, \qquad \big[g^{\alpha}_{\ \beta';\gamma\delta'}\big] = \frac{1}{2}\,R^{\alpha'}_{\ \beta'\gamma'\delta'}, \\
\big[g^{\alpha}_{\ \beta';\gamma'\delta}\big] &= -\frac{1}{2}\,R^{\alpha'}_{\ \beta'\gamma'\delta'}, \qquad \big[g^{\alpha}_{\ \beta';\gamma'\delta'}\big] = \frac{1}{2}\,R^{\alpha'}_{\ \beta'\gamma'\delta'}.
\end{aligned} \tag{5.15}$$

6 Expansion of bitensors near coincidence

6.1 General method

We would like to express a bitensor $\Omega_{\alpha'\beta'}(x,x')$ near coincidence as an expansion in powers of $-\sigma^{\alpha'}(x,x')$, the closest analogue in curved spacetime to the flat-spacetime quantity $(x-x')^{\alpha}$. For concreteness we shall consider the case of rank-2 bitensor, and for the moment we will assume that the tensorial indices all refer to the base point x'.

The expansion we seek is of the form

$$\Omega_{\alpha'\beta'}(x,x') = A_{\alpha'\beta'} + A_{\alpha'\beta'\gamma'}\,\sigma^{\gamma'} + \frac{1}{2}\,A_{\alpha'\beta'\gamma'\delta'}\,\sigma^{\gamma'}\sigma^{\delta'} + O(\epsilon^{3}), \tag{6.1}$$

in which the "expansion coefficients" $A_{\alpha'\beta'}$, $A_{\alpha'\beta'\gamma'}$, and $A_{\alpha'\beta'\gamma'\delta'}$ are all ordinary tensors at x'; this last tensor is symmetric in the pair of indices γ' and δ', and ϵ measures the size of a typical component of $\sigma^{\alpha'}$.

To find the expansion coefficients we differentiate Eq. (6.1) repeatedly and take coincidence limits. Equation (6.1) immediately implies $[\Omega_{\alpha'\beta'}] = A_{\alpha'\beta'}$. After one differentiation we obtain $\Omega_{\alpha'\beta';\gamma'} = A_{\alpha'\beta';\gamma'} + A_{\alpha'\beta'\epsilon';\gamma'}\sigma^{\epsilon'} + A_{\alpha'\beta'\epsilon'}\sigma^{\epsilon'}{}_{\gamma'} + \frac{1}{2}A_{\alpha'\beta'\epsilon'\iota';\gamma'}\sigma^{\epsilon'}\sigma^{\iota'} + A_{\alpha'\beta'\epsilon'\iota'}\sigma^{\epsilon'}\sigma^{\iota'}{}_{\gamma'} + O(\epsilon^2)$, and at coincidence this reduces to $[\Omega_{\alpha'\beta';\gamma'}] = A_{\alpha'\beta';\gamma'} + A_{\alpha'\beta'\gamma'}$. Taking the coincidence limit after two differentiations yields $[\Omega_{\alpha'\beta';\gamma'\delta'}] = A_{\alpha'\beta';\gamma'\delta'} + A_{\alpha'\beta'\gamma';\delta'} + A_{\alpha'\beta'\delta';\gamma'} + A_{\alpha'\beta'\gamma'\delta'}$. The expansion coefficients are therefore

$$\begin{aligned}
A_{\alpha'\beta'} &= \big[\Omega_{\alpha'\beta'}\big], \\
A_{\alpha'\beta'\gamma'} &= \big[\Omega_{\alpha'\beta';\gamma'}\big] - A_{\alpha'\beta';\gamma'}, \\
A_{\alpha'\beta'\gamma'\delta'} &= \big[\Omega_{\alpha'\beta';\gamma'\delta'}\big] - A_{\alpha'\beta';\gamma'\delta'} - A_{\alpha'\beta'\gamma';\delta'} - A_{\alpha'\beta'\delta';\gamma'}.
\end{aligned} \tag{6.2}$$

These results are to be substituted into Eq. (6.1), and this gives us $\Omega_{\alpha'\beta'}(x,x')$ to second order in ϵ.

Suppose now that the bitensor is $\Omega_{\alpha'\beta}$, with one index referring to x' and the other to x. The previous procedure can be applied directly if we introduce an auxiliary bitensor $\tilde{\Omega}_{\alpha'\beta'} := g^{\beta}_{\ \beta'}\Omega_{\alpha'\beta}$ whose indices all refer to the point x'. Then $\tilde{\Omega}_{\alpha'\beta'}$ can be expanded as in Eq. (6.1), and the original bitensor is reconstructed as $\Omega_{\alpha'\beta} = g^{\beta'}_{\ \beta}\tilde{\Omega}_{\alpha'\beta'}$, or

$$\Omega_{\alpha'\beta}(x,x') = g^{\beta'}_{\ \beta}\bigg(B_{\alpha'\beta'} + B_{\alpha'\beta'\gamma'}\,\sigma^{\gamma'} + \frac{1}{2}\,B_{\alpha'\beta'\gamma'\delta'}\,\sigma^{\gamma'}\sigma^{\delta'}\bigg) + O(\epsilon^3). \tag{6.3}$$

The expansion coefficients can be obtained from the coincidence limits of $\tilde{\Omega}_{\alpha'\beta'}$ and its derivatives. It is convenient, however, to express them directly in terms of the original bitensor $\Omega_{\alpha'\beta}$ by substituting the relation $\tilde{\Omega}_{\alpha'\beta'} = g^{\beta}_{\ \beta'}\Omega_{\alpha'\beta}$ and its derivatives. After using the results of Eq. (5.13)–(5.15) we find

$$\begin{aligned}
B_{\alpha'\beta'} &= \big[\Omega_{\alpha'\beta}\big], \\
B_{\alpha'\beta'\gamma'} &= \big[\Omega_{\alpha'\beta;\gamma'}\big] - B_{\alpha'\beta';\gamma'}, \\
B_{\alpha'\beta'\gamma'\delta'} &= \big[\Omega_{\alpha'\beta;\gamma'\delta'}\big] + \frac{1}{2}\,B_{\alpha'\epsilon'}R^{\epsilon'}_{\ \beta'\gamma'\delta'} - B_{\alpha'\beta';\gamma'\delta'} - B_{\alpha'\beta'\gamma';\delta'} - B_{\alpha'\beta'\delta';\gamma'}.
\end{aligned} \tag{6.4}$$

The only difference with respect to Eq. (6.3) is the presence of a Riemann-tensor term in $B_{\alpha'\beta'\gamma'\delta'}$.

Suppose finally that the bitensor to be expanded is $\Omega_{\alpha\beta}$, whose indices all refer to x. Much as we did before, we introduce an auxiliary bitensor $\tilde{\Omega}_{\alpha'\beta'} = g^{\alpha}_{\ \alpha'}g^{\beta}_{\ \beta'}\Omega_{\alpha\beta}$ whose indices all refer to x', we expand $\tilde{\Omega}_{\alpha'\beta'}$ as in Eq. (6.1), and we then reconstruct the original bitensor. This gives us

$$\Omega_{\alpha\beta}(x,x') = g^{\alpha'}_{\ \alpha}g^{\beta'}_{\ \beta}\bigg(C_{\alpha'\beta'} + C_{\alpha'\beta'\gamma'}\,\sigma^{\gamma'} + \frac{1}{2}\,C_{\alpha'\beta'\gamma'\delta'}\,\sigma^{\gamma'}\sigma^{\delta'}\bigg) + O(\epsilon^3), \tag{6.5}$$

and the expansion coefficients are now

$$\begin{aligned}
C_{\alpha'\beta'} &= \big[\Omega_{\alpha\beta}\big], \\
C_{\alpha'\beta'\gamma'} &= \big[\Omega_{\alpha\beta;\gamma'}\big] - C_{\alpha'\beta';\gamma'}, \\
C_{\alpha'\beta'\gamma'\delta'} &= \big[\Omega_{\alpha\beta;\gamma'\delta'}\big] + \frac{1}{2}\,C_{\alpha'\epsilon'}R^{\epsilon'}_{\ \beta'\gamma'\delta'} + \frac{1}{2}\,C_{\epsilon'\beta'}R^{\epsilon'}_{\ \alpha'\gamma'\delta'} - C_{\alpha'\beta';\gamma'\delta'} - C_{\alpha'\beta'\gamma';\delta'} - C_{\alpha'\beta'\delta';\gamma'}.
\end{aligned} \tag{6.6}$$

This differs from Eq. (6.4) by the presence of an additional Riemann-tensor term in $C_{\alpha'\beta'\gamma'\delta'}$.

6.2 Special cases

We now apply the general expansion method developed in the preceding subsection to the bitensors $\sigma_{\alpha'\beta'}$, $\sigma_{\alpha'\beta}$, and $\sigma_{\alpha\beta}$. In the first instance we have $A_{\alpha'\beta'} = g_{\alpha'\beta'}$, $A_{\alpha'\beta'\gamma'} = 0$, and $A_{\alpha'\beta'\gamma'\delta'} = -\frac{1}{3}(R_{\alpha'\gamma'\beta'\delta'} + R_{\alpha'\delta'\beta'\gamma'})$. In the second instance we have $B_{\alpha'\beta'} = -g_{\alpha'\beta'}$, $B_{\alpha'\beta'\gamma'} = 0$, and $B_{\alpha'\beta'\gamma'\delta'} = -\frac{1}{3}(R_{\beta'\alpha'\gamma'\delta'} +$

$R_{\beta'\gamma'\alpha'\delta'}) - \frac{1}{2}R_{\alpha'\beta'\gamma'\delta'} = -\frac{1}{3}R_{\alpha'\delta'\beta'\gamma'} - \frac{1}{6}R_{\alpha'\beta'\gamma'\delta'}$. In the third instance we have $C_{\alpha'\beta'} = g_{\alpha'\beta'}$, $C_{\alpha'\beta'\gamma'} = 0$, and $C_{\alpha'\beta'\gamma'\delta'} = -\frac{1}{3}(R_{\alpha'\gamma'\beta'\delta'} + R_{\alpha'\delta'\beta'\gamma'})$. This gives us the expansions

$$\sigma_{\alpha'\beta'} = g_{\alpha'\beta'} - \frac{1}{3}\, R_{\alpha'\gamma'\beta'\delta'}\, \sigma^{\gamma'}\sigma^{\delta'} + O(\epsilon^3), \tag{6.7}$$

$$\sigma_{\alpha'\beta} = -g^{\beta'}_{\ \beta}\Big(g_{\alpha'\beta'} + \frac{1}{6}\, R_{\alpha'\gamma'\beta'\delta'}\, \sigma^{\gamma'}\sigma^{\delta'}\Big) + O(\epsilon^3), \tag{6.8}$$

$$\sigma_{\alpha\beta} = g^{\alpha'}_{\ \alpha} g^{\beta'}_{\ \beta'}\Big(g_{\alpha'\beta'} - \frac{1}{3}\, R_{\alpha'\gamma'\beta'\delta'}\, \sigma^{\gamma'}\sigma^{\delta'}\Big) + O(\epsilon^3). \tag{6.9}$$

Taking the trace of the last equation returns $\sigma^{\alpha}_{\ \alpha} = 4 - \frac{1}{3}R_{\gamma'\delta'}\,\sigma^{\gamma'}\sigma^{\delta'} + O(\epsilon^3)$, or

$$\theta^* = 3 - \frac{1}{3}\, R_{\alpha'\beta'}\, \sigma^{\alpha'}\sigma^{\beta'} + O(\epsilon^3), \tag{6.10}$$

where $\theta^* := \sigma^{\alpha}_{\ \alpha} - 1$ was shown in Sec. 3.4 to describe the expansion of the congruence of geodesics that emanate from x'. Equation (6.10) reveals that timelike geodesics are focused if the Ricci tensor is nonzero and the strong energy condition holds: when $R_{\alpha'\beta'}\,\sigma^{\alpha'}\sigma^{\beta'} > 0$ we see that θ^* is smaller than 3, the value it would take in flat spacetime.

The expansion method can easily be extended to bitensors of other tensorial ranks. In particular, it can be adapted to give expansions of the first derivatives of the parallel propagator. The expansions

$$g^{\alpha}_{\ \beta';\gamma'} = \frac{1}{2}\, g^{\alpha}_{\ \alpha'} R^{\alpha'}_{\ \beta'\gamma'\delta'}\, \sigma^{\delta'} + O(\epsilon^2), \qquad g^{\alpha}_{\ \beta';\gamma} = \frac{1}{2}\, g^{\alpha}_{\ \alpha'} g^{\gamma'}_{\ \gamma} R^{\alpha'}_{\ \beta'\gamma'\delta'}\, \sigma^{\delta'} + O(\epsilon^2) \tag{6.11}$$

and thus easy to establish, and they will be needed in part III of this review.

6.3 Expansion of tensors

The expansion method can also be applied to ordinary tensor fields. For concreteness, suppose that we wish to express a rank-2 tensor $A_{\alpha\beta}$ at a point x in terms of its values (and that of its covariant derivatives) at a neighbouring point x'. The tensor can be written as an expansion in powers of $-\sigma^{\alpha'}(x, x')$ and in this case we have

$$A_{\alpha\beta}(x) = g^{\alpha'}_{\ \alpha} g^{\beta'}_{\ \beta}\bigg(A_{\alpha'\beta'} - A_{\alpha'\beta';\gamma'}\, \sigma^{\gamma'} + \frac{1}{2}\, A_{\alpha'\beta';\gamma'\delta'}\, \sigma^{\gamma'}\sigma^{\delta'}\bigg) + O(\epsilon^3). \tag{6.12}$$

If the tensor field is parallel transported on the geodesic β that links x to x', then Eq. (6.12) reduces to Eq. (5.10). The extension of this formula to tensors of other ranks is obvious.

To derive this result we express $A_{\mu\nu}(z)$, the restriction of the tensor field on β, in terms of its tetrad components $A_{\mathsf{ab}}(\lambda) = A_{\mu\nu} e^{\mu}_{\mathsf{a}} e^{\nu}_{\mathsf{b}}$. Recall from Sec. 5.1 that e^{μ}_{a} is an orthonormal basis that is parallel transported on β; recall also that the affine parameter λ ranges from λ_0 (its value at x') to λ_1 (its value at x). We have $A_{\alpha'\beta'}(x') = A_{\mathsf{ab}}(\lambda_0) e^{\mathsf{a}}_{\alpha'} e^{\mathsf{b}}_{\beta'}$, $A_{\alpha\beta}(x) = A_{\mathsf{ab}}(\lambda_1) e^{\mathsf{a}}_{\alpha} e^{\mathsf{b}}_{\beta}$, and $A_{\mathsf{ab}}(\lambda_1)$ can be expressed in terms of quantities at $\lambda = \lambda_0$ by straightforward Taylor expansion. Since, for example,

$$(\lambda_1 - \lambda_0)\frac{dA_{\mathsf{ab}}}{d\lambda}\bigg|_{\lambda_0} = (\lambda_1 - \lambda_0)\big(A_{\mu\nu} e^{\mu}_{\mathsf{a}} e^{\nu}_{\mathsf{b}}\big)_{;\lambda} t^{\lambda}\Big|_{\lambda_0} = (\lambda_1 - \lambda_0) A_{\mu\nu;\lambda} e^{\mu}_{\mathsf{a}} e^{\nu}_{\mathsf{b}} t^{\lambda}\Big|_{\lambda_0} = -A_{\alpha'\beta';\gamma'} e^{\alpha'}_{\mathsf{a}} e^{\beta'}_{\mathsf{b}} \sigma^{\gamma'},$$

where we have used Eq. (3.4), we arrive at Eq. (6.12) after involving Eq. (5.6).

7 van Vleck determinant

7.1 Definition and properties

The van Vleck biscalar $\Delta(x, x')$ is defined by

$$\Delta(x, x') := \det\big[\Delta^{\alpha'}_{\ \beta'}(x, x')\big], \qquad \Delta^{\alpha'}_{\ \beta'}(x, x') := -g^{\alpha'}_{\ \alpha}(x', x)\sigma^{\alpha}_{\ \beta'}(x, x'). \tag{7.1}$$

As we shall show below, it can also be expressed as

$$\Delta(x, x') = -\frac{\det\bigl[-\sigma_{\alpha\beta'}(x, x')\bigr]}{\sqrt{-g}\sqrt{-g'}}, \tag{7.2}$$

where g is the metric determinant at x and g' the metric determinant at x'.

Equations (4.2) and (5.13) imply that at coincidence, $[\Delta^{\alpha'}{}_{\beta'}] = \delta^{\alpha'}{}_{\beta'}$ and $[\Delta] = 1$. Equation (6.8), on the other hand, implies that near coincidence,

$$\Delta^{\alpha'}{}_{\beta'} = \delta^{\alpha'}{}_{\beta'} + \frac{1}{6} R^{\alpha'}{}_{\gamma'\beta'\delta'}\, \sigma^{\gamma'} \sigma^{\delta'} + O(\epsilon^3), \tag{7.3}$$

so that

$$\Delta = 1 + \frac{1}{6} R_{\alpha'\beta'}\, \sigma^{\alpha'} \sigma^{\beta'} + O(\epsilon^3). \tag{7.4}$$

This last result follows from the fact that for a "small" matrix $\boldsymbol{a}$, $\det(\boldsymbol{1} + \boldsymbol{a}) = 1 + \mathrm{tr}(\boldsymbol{a}) + O(\boldsymbol{a}^2)$.

We shall prove below that the van Vleck determinant satisfies the differential equation

$$\frac{1}{\Delta}\bigl(\Delta\sigma^\alpha\bigr)_{;\alpha} = 4, \tag{7.5}$$

which can also be written as $(\ln\Delta)_{,\alpha}\sigma^\alpha = 4 - \sigma^\alpha{}_\alpha$, or

$$\frac{d}{d\lambda^*}(\ln\Delta) = 3 - \theta^* \tag{7.6}$$

in the notation introduced in Sec. 3.4. Equation (7.6) reveals that the behaviour of the van Vleck determinant is governed by the expansion of the congruence of geodesics that emanate from x'. If $\theta^* < 3$, then the congruence expands less rapidly than it would in flat spacetime, and Δ *increases* along the geodesics. If, on the other hand, $\theta^* > 3$, then the congruence expands more rapidly than it would in flat spacetime, and Δ *decreases* along the geodesics. Thus, $\Delta > 1$ indicates that the geodesics are undergoing focusing, while $\Delta < 1$ indicates that the geodesics are undergoing defocusing. The connection between the van Vleck determinant and the strong energy condition is well illustrated by Eq. (7.4): the sign of $\Delta - 1$ near x' is determined by the sign of $R_{\alpha'\beta'}\, \sigma^{\alpha'}\sigma^{\beta'}$.

7.2 Derivations

To show that Eq. (7.2) follows from Eq. (7.1) we rewrite the completeness relations at x, $g^{\alpha\beta} = \eta^{\mathsf{ab}} e^\alpha_{\mathsf{a}} e^\beta_{\mathsf{b}}$, in the matrix form $\boldsymbol{g}^{-1} = \boldsymbol{E}\boldsymbol{\eta}\boldsymbol{E}^T$, where $\boldsymbol{E}$ denotes the 4×4 matrix whose entries correspond to e^α_{a}. (In this translation we put tensor and frame indices on an equal footing.) With e denoting the determinant of this matrix, we have $1/g = -e^2$, or $e = 1/\sqrt{-g}$. Similarly, we rewrite the completeness relations at x', $g^{\alpha'\beta'} = \eta^{\mathsf{ab}} e^{\alpha'}_{\mathsf{a}} e^{\beta'}_{\mathsf{b}}$, in the matrix form $\boldsymbol{g'}^{-1} = \boldsymbol{E'}\boldsymbol{\eta}\boldsymbol{E'}^T$, where $\boldsymbol{E'}$ is the matrix corresponding to $e^{\alpha'}_{\mathsf{a}}$. With e' denoting its determinant, we have $1/g' = -e'^2$, or $e' = 1/\sqrt{-g'}$. Now, the parallel propagator is defined by $g^\alpha{}_{\alpha'} = \eta^{\mathsf{ab}} g_{\alpha'\beta'} e^\alpha_{\mathsf{a}} e^{\beta'}_{\mathsf{b}}$, and the matrix form of this equation is $\hat{\boldsymbol{g}} = \boldsymbol{E}\boldsymbol{\eta}\boldsymbol{E'}^T\boldsymbol{g'}^T$. The determinant of the parallel propagator is therefore $\hat{g} = -ee'g' = \sqrt{-g'}/\sqrt{-g}$. So we have

$$\det\bigl[g^\alpha{}_{\alpha'}\bigr] = \frac{\sqrt{-g'}}{\sqrt{-g}}, \qquad \det\bigl[g^{\alpha'}{}_\alpha\bigr] = \frac{\sqrt{-g}}{\sqrt{-g'}}, \tag{7.7}$$

and Eq. (7.2) follows from the fact that the matrix form of Eq. (7.1) is $\boldsymbol{\Delta} = -\hat{\boldsymbol{g}}^{-1}\boldsymbol{g}^{-1}\boldsymbol{\sigma}$, where $\boldsymbol{\sigma}$ is the matrix corresponding to $\sigma_{\alpha\beta'}$.

To establish Eq. (7.5) we differentiate the relation $\sigma = \frac{1}{2}\sigma^\gamma\sigma_\gamma$ twice and obtain $\sigma_{\alpha\beta'} = \sigma^\gamma{}_\alpha\sigma_{\gamma\beta'} + \sigma^\gamma\sigma_{\gamma\alpha\beta'}$. If we replace the last factor by $\sigma_{\alpha\beta'\gamma}$ and multiply both sides by $-g^{\alpha'\alpha}$ we find

$$\Delta^{\alpha'}{}_{\beta'} = -g^{\alpha'\alpha}\bigl(\sigma^\gamma{}_\alpha\sigma_{\gamma\beta'} + \sigma^\gamma\sigma_{\alpha\beta'\gamma}\bigr).$$

In this expression we make the substitution $\sigma_{\alpha\beta'} = -g_{\alpha\alpha'}\Delta^{\alpha'}{}_{\beta'}$, which follows directly from Eq. (7.1). This gives us

$$\Delta^{\alpha'}{}_{\beta'} = g^{\alpha'}{}_{\alpha} g^{\gamma}{}_{\gamma'} \sigma^{\alpha}{}_{\gamma} \Delta^{\gamma'}{}_{\beta'} + \Delta^{\alpha'}{}_{\beta';\gamma}\sigma^{\gamma}, \tag{7.8}$$

where we have used Eq. (5.11). At this stage we introduce an inverse $(\Delta^{-1})^{\alpha'}{}_{\beta'}$ to the van Vleck bitensor, defined by $\Delta^{\alpha'}{}_{\beta'}(\Delta^{-1})^{\beta'}{}_{\gamma'} = \delta^{\alpha'}{}_{\gamma'}$. After multiplying both sides of Eq. (7.8) by $(\Delta^{-1})^{\beta'}{}_{\gamma'}$ we find

$$\delta^{\alpha'}{}_{\beta'} = g^{\alpha'}{}_{\alpha} g^{\beta}{}_{\beta'} \sigma^{\alpha}{}_{\beta} + (\Delta^{-1})^{\gamma'}{}_{\beta'} \Delta^{\alpha'}{}_{\gamma';\gamma}\sigma^{\gamma},$$

and taking the trace of this equation yields

$$4 = \sigma^{\alpha}{}_{\alpha} + (\Delta^{-1})^{\beta'}{}_{\alpha'} \Delta^{\alpha'}{}_{\beta';\gamma}\sigma^{\gamma}.$$

We now recall the identity $\delta \ln \det \boldsymbol{M} = \mathrm{Tr}(\boldsymbol{M}^{-1}\delta\boldsymbol{M})$, which relates the variation of a determinant to the variation of the matrix elements. It implies, in particular, that $(\Delta^{-1})^{\beta'}{}_{\alpha'} \Delta^{\alpha'}{}_{\beta';\gamma} = (\ln\Delta)_{,\gamma}$, and we finally obtain

$$4 = \sigma^{\alpha}{}_{\alpha} + (\ln\Delta)_{,\alpha}\sigma^{\alpha}, \tag{7.9}$$

which is equivalent to Eq. (7.5) or Eq. (7.6).

Part II
Coordinate systems

8 Riemann normal coordinates

8.1 Definition and coordinate transformation

Given a fixed base point x' and a tetrad $e^{\alpha'}_{\mathsf{a}}(x')$, we assign to a neighbouring point x the four coordinates

$$\hat{x}^{\mathsf{a}} = -e^{\mathsf{a}}_{\alpha'}(x')\,\sigma^{\alpha'}(x,x'), \tag{8.1}$$

where $e^{\mathsf{a}}_{\alpha'} = \eta^{\mathsf{ab}} g_{\alpha'\beta'} e^{\beta'}_{\mathsf{b}}$ is the dual tetrad attached to x'. The new coordinates $\hat{x}^{\mathsf{a}}$ are called *Riemann normal coordinates* (RNC), and they are such that $\eta_{\mathsf{ab}}\hat{x}^{\mathsf{a}}\hat{x}^{\mathsf{b}} = \eta_{\mathsf{ab}} e^{\mathsf{a}}_{\alpha'} e^{\mathsf{b}}_{\beta'} \sigma^{\alpha'}\sigma^{\beta'} = g_{\alpha'\beta'}\sigma^{\alpha'}\sigma^{\beta'}$, or

$$\eta_{\mathsf{ab}}\hat{x}^{\mathsf{a}}\hat{x}^{\mathsf{b}} = 2\sigma(x,x'). \tag{8.2}$$

Thus, $\eta_{\mathsf{ab}}\hat{x}^{\mathsf{a}}\hat{x}^{\mathsf{b}}$ is the squared geodesic distance between x and the base point x'. It is obvious that x' is at the origin of the RNC, where $\hat{x}^{\mathsf{a}} = 0$.

If we move the point x to $x + \delta x$, the new coordinates change to $\hat{x}^{\mathsf{a}} + \delta\hat{x}^{\mathsf{a}} = -e^{\mathsf{a}}_{\alpha'}\sigma^{\alpha'}(x+\delta x, x') = \hat{x}^{\mathsf{a}} - e^{\mathsf{a}}_{\alpha'}\sigma^{\alpha'}_{\ \ \beta}\,\delta x^{\beta}$, so that

$$d\hat{x}^{\mathsf{a}} = -e^{\mathsf{a}}_{\alpha'}\sigma^{\alpha'}_{\ \ \beta}\,dx^{\beta}. \tag{8.3}$$

The coordinate transformation is therefore determined by $\partial\hat{x}^{\mathsf{a}}/\partial x^{\beta} = -e^{\mathsf{a}}_{\alpha'}\sigma^{\alpha'}_{\ \ \beta}$, and at coincidence we have

$$\left[\frac{\partial\hat{x}^{\mathsf{a}}}{\partial x^{\alpha}}\right] = e^{\mathsf{a}}_{\alpha'}, \qquad \left[\frac{\partial x^{\alpha}}{\partial\hat{x}^{\mathsf{a}}}\right] = e^{\alpha'}_{\mathsf{a}}; \tag{8.4}$$

the second result follows from the identities $e^{\mathsf{a}}_{\alpha'}e^{\alpha'}_{\mathsf{b}} = \delta^{\mathsf{a}}_{\ \mathsf{b}}$ and $e^{\alpha'}_{\mathsf{a}}e^{\mathsf{a}}_{\beta'} = \delta^{\alpha'}_{\ \beta'}$.

It is interesting to note that the Jacobian of the transformation of Eq. (8.3), $J := \det(\partial\hat{x}^{\mathsf{a}}/\partial x^{\beta})$, is given by $J = \sqrt{-g}\Delta(x,x')$, where g is the determinant of the metric in the original coordinates, and $\Delta(x,x')$ is the Van Vleck determinant of Eq. (7.2). This result follows simply by writing the coordinate transformation in the form $\partial\hat{x}^{\mathsf{a}}/\partial x^{\beta} = -\eta^{\mathsf{ab}}e^{\alpha'}_{\mathsf{b}}\sigma_{\alpha'\beta}$ and computing the product of the determinants. It allows us to deduce that in RNC, the determinant of the metric is given by

$$\sqrt{-g(\mathrm{RNC})} = \frac{1}{\Delta(x,x')}. \tag{8.5}$$

It is easy to show that the geodesics emanating from x' are straight coordinate lines in RNC. The proper volume of a small comoving region is then equal to $dV = \Delta^{-1}\,d^4\hat{x}$, and this is smaller than the flat-spacetime value of $d^4\hat{x}$ if $\Delta > 1$, that is, if the geodesics are focused by the spacetime curvature.

8.2 Metric near x'

We now would like to invert Eq. (8.3) in order to express the line element $ds^2 = g_{\alpha\beta}\,dx^{\alpha}dx^{\beta}$ in terms of the displacements $d\hat{x}^{\mathsf{a}}$. We shall do this approximately, by working in a small neighbourhood of x'. We recall the expansion of Eq. (6.8),

$$\sigma^{\alpha'}_{\ \ \beta} = -g^{\beta'}_{\ \ \beta}\left(\delta^{\alpha'}_{\ \ \beta'} + \frac{1}{6}R^{\alpha'}_{\ \ \gamma'\beta'\delta'}\sigma^{\gamma'}\sigma^{\delta'}\right) + O(\epsilon^3),$$

and in this we substitute the frame decomposition of the Riemann tensor, $R^{\alpha'}{}_{\gamma'\beta'\delta'} = R^{\mathsf{a}}{}_{\mathsf{cbd}}\, e^{\alpha'}_{\mathsf{a}} e^{\mathsf{c}}_{\gamma'} e^{\mathsf{b}}_{\beta'} e^{\mathsf{d}}_{\delta'}$, and the tetrad decomposition of the parallel propagator, $g^{\beta'}{}_{\beta} = e^{\beta'}_{\mathsf{b}}\, e^{\mathsf{b}}_{\beta}$, where $e^{\mathsf{b}}_{\beta}(x)$ is the dual tetrad at x obtained by parallel transport of $e^{\mathsf{b}}_{\beta'}(x')$. After some algebra we obtain

$$\sigma^{\alpha'}{}_{\beta} = -e^{\alpha'}_{\mathsf{a}}\, e^{\mathsf{a}}_{\beta} - \frac{1}{6}\, R^{\mathsf{a}}{}_{\mathsf{cbd}}\, e^{\alpha'}_{\mathsf{a}}\, e^{\mathsf{b}}_{\beta} \hat{x}^{\mathsf{c}} \hat{x}^{\mathsf{d}} + O(\epsilon^3),$$

where we have used Eq. (8.1). Substituting this into Eq. (8.3) yields

$$d\hat{x}^{\mathsf{a}} = \biggl[\delta^{\mathsf{a}}{}_{\mathsf{b}} + \frac{1}{6}\, R^{\mathsf{a}}{}_{\mathsf{cbd}} \hat{x}^{\mathsf{c}} \hat{x}^{\mathsf{d}} + O(x^3) \biggr] e^{\mathsf{b}}_{\beta}\, dx^{\beta}, \tag{8.6}$$

and this is easily inverted to give

$$e^{\mathsf{a}}_{\alpha}\, dx^{\alpha} = \biggl[\delta^{\mathsf{a}}{}_{\mathsf{b}} - \frac{1}{6}\, R^{\mathsf{a}}{}_{\mathsf{cbd}} \hat{x}^{\mathsf{c}} \hat{x}^{\mathsf{d}} + O(x^3) \biggr] d\hat{x}^{\mathsf{b}}. \tag{8.7}$$

This is the desired approximate inversion of Eq. (8.3). It is useful to note that Eq. (8.7), when specialized from the arbitrary coordinates x^α to $\hat{x}^{\mathsf{a}}$, gives us the components of the dual tetrad at x in RNC. And since $e^{\alpha'}_{\mathsf{a}} = \delta^{\alpha'}_{\mathsf{a}}$ in RNC, we immediately obtain the components of the parallel propagator: $g^{\mathsf{a}'}{}_{\mathsf{b}} = \delta^{\mathsf{a}}{}_{\mathsf{b}} - \frac{1}{6} R^{\mathsf{a}}{}_{\mathsf{cbd}} \hat{x}^{\mathsf{c}} \hat{x}^{\mathsf{d}} + O(x^3)$.

We are now in a position to calculate the metric in the new coordinates. We have $ds^2 = g_{\alpha\beta}\, dx^\alpha dx^\beta = (\eta_{\mathsf{ab}} e^{\mathsf{a}}_{\alpha} e^{\mathsf{b}}_{\beta}) dx^\alpha dx^\beta = \eta_{\mathsf{ab}} (e^{\mathsf{a}}_{\alpha}\, dx^\alpha)(e^{\mathsf{b}}_{\beta}\, dx^\beta)$, and in this we substitute Eq. (8.7). The final result is $ds^2 = g_{\mathsf{ab}}\, d\hat{x}^{\mathsf{a}} d\hat{x}^{\mathsf{b}}$, with

$$g_{\mathsf{ab}} = \eta_{\mathsf{ab}} - \frac{1}{3} R_{\mathsf{acbd}} \hat{x}^{\mathsf{c}} \hat{x}^{\mathsf{d}} + O(x^3). \tag{8.8}$$

The quantities R_{acbd} appearing in Eq. (8.8) are the frame components of the Riemann tensor evaluated at the base point x',

$$R_{\mathsf{acbd}} := R_{\alpha'\gamma'\beta'\delta'}\, e^{\alpha'}_{\mathsf{a}}\, e^{\gamma'}_{\mathsf{c}}\, e^{\beta'}_{\mathsf{b}}\, e^{\delta'}_{\mathsf{d}}, \tag{8.9}$$

and these are independent of $\hat{x}^{\mathsf{a}}$. They are also, by virtue of Eq. (8.4), the components of the (base-point) Riemann tensor in RNC, because Eq. (8.9) can also be expressed as

$$R_{\mathsf{acdb}} = R_{\alpha'\gamma'\beta'\delta'} \biggl[\frac{\partial x^\alpha}{\partial \hat{x}^{\mathsf{a}}} \biggr] \biggl[\frac{\partial x^\gamma}{\partial \hat{x}^{\mathsf{c}}} \biggr] \biggl[\frac{\partial x^\beta}{\partial \hat{x}^{\mathsf{b}}} \biggr] \biggl[\frac{\partial x^\delta}{\partial \hat{x}^{\mathsf{d}}} \biggr],$$

which is the standard transformation law for tensor components.

It is obvious from Eq. (8.8) that $g_{\mathsf{ab}}(x') = \eta_{\mathsf{ab}}$ and $\Gamma^{\mathsf{a}}{}_{\mathsf{bc}}(x') = 0$, where $\Gamma^{\mathsf{a}}{}_{\mathsf{bc}} = -\frac{1}{3}(R^{\mathsf{a}}{}_{\mathsf{bcd}} + R^{\mathsf{a}}{}_{\mathsf{cbd}})\hat{x}^{\mathsf{d}} + O(x^2)$ is the connection compatible with the metric g_{ab}. The Riemann normal coordinates therefore provide a constructive proof of the local flatness theorem.

9 Fermi normal coordinates

9.1 Fermi-Walker transport

Let γ be a timelike curve described by parametric relations $z^\mu(\tau)$ in which τ is proper time. Let $u^\mu = dz^\mu/d\tau$ be the curve's normalized tangent vector, and let $a^\mu = Du^\mu/d\tau$ be its acceleration vector.

A vector field v^μ is said to be *Fermi-Walker transported* on γ if it is a solution to the differential equation

$$\frac{Dv^\mu}{d\tau} = \bigl(v_\nu a^\nu\bigr) u^\mu - \bigl(v_\nu u^\nu\bigr) a^\mu. \tag{9.1}$$

Notice that this reduces to parallel transport when $a^\mu = 0$ and γ is a geodesic.

The operation of Fermi-Walker (FW) transport satisfies two important properties. The first is that u^μ is automatically FW transported along γ; this follows at once from Eq. (9.1) and the fact that u^μ is orthogonal to a^μ. The second is that if the vectors v^μ and w^μ are both FW transported along γ, then their inner product $v_\mu w^\mu$ is constant on γ: $D(v_\mu w^\mu)/d\tau = 0$; this also follows immediately from Eq. (9.1).

9.2 Tetrad and dual tetrad on γ

Let $\bar{z}$ be an arbitrary reference point on γ. At this point we erect an orthonormal tetrad $(u^{\bar{\mu}}, e^{\bar{\mu}}_a)$ where, as a modification to former usage, the frame index a runs from 1 to 3. We then propagate each frame vector on γ by FW transport; this guarantees that the tetrad remains orthonormal everywhere on γ. At a generic point $z(\tau)$ we have

$$\frac{De^{\mu}_a}{d\tau} = \bigl(a_\nu e^\nu_a\bigr)u^\mu, \qquad g_{\mu\nu}u^\mu u^\nu = -1, \qquad g_{\mu\nu}e^\mu_a u^\nu = 0, \qquad g_{\mu\nu}e^\mu_a e^\nu_b = \delta_{ab}. \tag{9.2}$$

From the tetrad on γ we define a dual tetrad (e^0_μ, e^a_μ) by the relations

$$e^0_\mu = -u_\mu, \qquad e^a_\mu = \delta^{ab} g_{\mu\nu} e^\nu_b; \tag{9.3}$$

this also is FW transported on γ. The tetrad and its dual give rise to the completeness relations

$$g^{\mu\nu} = -u^\mu u^\nu + \delta^{ab}e^\mu_a e^\nu_b, \qquad g_{\mu\nu} = -e^0_\mu e^0_\nu + \delta_{ab}\, e^a_\mu e^b_\nu. \tag{9.4}$$

9.3 Fermi normal coordinates

To construct the Fermi normal coordinates (FNC) of a point x in the normal convex neighbourhood of γ we locate the unique *spacelike geodesic* β that passes through x and intersects γ *orthogonally.* We denote the intersection point by $\bar{x} := z(t)$, with t denoting the value of the proper-time parameter at this point. To tensors at $\bar{x}$ we assign indices $\bar{\alpha}$, $\bar{\beta}$, and so on. The FNC of x are defined by

$$\hat{x}^0 = t, \qquad \hat{x}^a = -e^a_{\bar{\alpha}}(\bar{x})\sigma^{\bar{\alpha}}(x,\bar{x}), \qquad \sigma_{\bar{\alpha}}(x,\bar{x})u^{\bar{\alpha}}(\bar{x}) = 0; \tag{9.5}$$

the last statement determines $\bar{x}$ from the requirement that $-\sigma^{\bar{\alpha}}$, the vector tangent to β at $\bar{x}$, be orthogonal to $u^{\bar{\alpha}}$, the vector tangent to γ. From the definition of the FNC and the completeness relations of Eq. (9.4) it follows that

$$s^2 := \delta_{ab}\hat{x}^a\hat{x}^b = 2\sigma(x,\bar{x}), \tag{9.6}$$

so that s is the spatial distance between $\bar{x}$ and x along the geodesic β. This statement gives an immediate meaning to $\hat{x}^a$, the spatial Fermi normal coordinates, and the time coordinate $\hat{x}^0$ is simply proper time at the intersection point $\bar{x}$. The situation is illustrated in Fig. 6.

Suppose that x is moved to $x+\delta x$. This typically induces a change in the spacelike geodesic β, which moves to $\beta+\delta\beta$, and a corresponding change in the intersection point $\bar{x}$, which moves to $x'' := \bar{x}+\delta\bar{x}$, with $\delta x^{\bar{\alpha}} = u^{\bar{\alpha}}\delta t$. The FNC of the new point are then $\hat{x}^0(x+\delta x) = t+\delta t$ and $\hat{x}^a(x+\delta x) = -e^a_{\alpha''}(x'')\sigma^{\alpha''}(x+\delta x, x'')$, with x'' determined by $\sigma_{\alpha''}(x+\delta x, x'')u^{\alpha''}(x'') = 0$. Expanding these relations to first order in the displacements, and simplifying using Eqs. (9.2), yields

$$dt = \mu\, \sigma_{\bar{\alpha}\beta}u^{\bar{\alpha}}\, dx^\beta, \qquad d\hat{x}^a = -e^a_{\bar{\alpha}}\bigl(\sigma^{\bar{\alpha}}{}_\beta + \mu\, \sigma^{\bar{\alpha}}{}_{\bar{\beta}}u^{\bar{\beta}}\sigma_{\beta\bar{\gamma}}u^{\bar{\gamma}}\bigr)dx^\beta, \tag{9.7}$$

where μ is determined by $\mu^{-1} = -(\sigma_{\bar{\alpha}\bar{\beta}}u^{\bar{\alpha}}u^{\bar{\beta}} + \sigma_{\bar{\alpha}}a^{\bar{\alpha}})$.

9.4 Coordinate displacements near γ

The relations of Eq. (9.7) can be expressed as expansions in powers of s, the spatial distance from $\bar{x}$ to x. For this we use the expansions of Eqs. (6.7) and (6.8), in which we substitute $\sigma^{\bar{\alpha}} = -e^{\bar{\alpha}}_a\hat{x}^a$ and $g^{\bar{\alpha}}{}_\alpha = u^{\bar{\alpha}}\bar{e}^0_\alpha + e^{\bar{\alpha}}_a\bar{e}^a_\alpha$, where $(\bar{e}^0_\alpha, \bar{e}^a_\alpha)$ is a dual tetrad at x obtained by parallel transport of $(-u_{\bar{\alpha}}, e^a_{\bar{\alpha}})$ on the spacelike geodesic β. After some algebra we obtain

$$\mu^{-1} = 1 + a_a\hat{x}^a + \frac{1}{3}R_{0c0d}\hat{x}^c\hat{x}^d + O(s^3),$$

where $a_a(t) := a_{\bar{\alpha}}e^{\bar{\alpha}}_a$ are frame components of the acceleration vector, and $R_{0c0d}(t) := R_{\bar{\alpha}\bar{\gamma}\bar{\beta}\bar{\delta}}u^{\bar{\alpha}}e^{\bar{\gamma}}_c u^{\bar{\beta}}e^{\bar{\delta}}_d$ are frame components of the Riemann tensor evaluated on γ. This last result is easily inverted to give

$$\mu = 1 - a_a\hat{x}^a + \bigl(a_a\hat{x}^a\bigr)^2 - \frac{1}{3}R_{0c0d}\hat{x}^c\hat{x}^d + O(s^3).$$

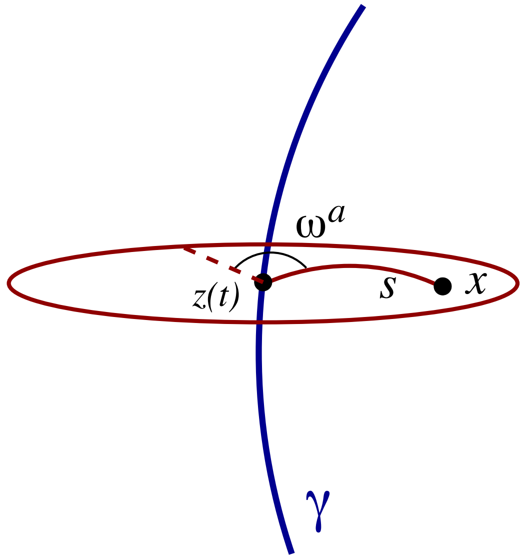

Figure 6: Fermi normal coordinates of a point x relative to a world line γ. The time coordinate t selects a particular point on the word line, and the disk represents the set of spacelike geodesics that intersect γ orthogonally at $z(t)$. The unit vector $\omega^a := \hat{x}^a/s$ selects a particular geodesic among this set, and the spatial distance s selects a particular point on this geodesic.

Proceeding similarly for the other relations of Eq. (9.7), we obtain

$$dt = \biggl[1 - a_a\hat{x}^a + \bigl(a_a\hat{x}^a\bigr)^2 - \frac{1}{2}R_{0c0d}\hat{x}^c\hat{x}^d + O(s^3)\biggr]\bigl(\bar{e}^0_\beta dx^\beta\bigr) + \biggl[-\frac{1}{6}R_{0cbd}\hat{x}^c\hat{x}^d + O(s^3)\biggr]\bigl(\bar{e}^b_\beta dx^\beta\bigr) \tag{9.8}$$

and

$$d\hat{x}^a = \biggl[\frac{1}{2}R^a{}_{c0d}\hat{x}^c\hat{x}^d + O(s^3)\biggr]\bigl(\bar{e}^0_\beta dx^\beta\bigr) + \biggl[\delta^a{}_b + \frac{1}{6}R^a{}_{cbd}\hat{x}^c\hat{x}^d + O(s^3)\biggr]\bigl(\bar{e}^b_\beta dx^\beta\bigr), \tag{9.9}$$

where $R_{ac0d}(t) := R_{\bar{\alpha}\bar{\gamma}\bar{\beta}\bar{\delta}}e^{\bar{\alpha}}_a e^{\bar{\gamma}}_c u^{\bar{\beta}} e^{\bar{\delta}}_d$ and $R_{acbd}(t) := R_{\bar{\alpha}\bar{\gamma}\bar{\beta}\bar{\delta}}e^{\bar{\alpha}}_a e^{\bar{\gamma}}_c e^{\bar{\beta}}_b e^{\bar{\delta}}_d$ are additional frame components of the Riemann tensor evaluated on γ. (Note that frame indices are raised with δ^{ab}.)

As a special case of Eqs. (9.8) and (9.9) we find that

$$\frac{\partial t}{\partial x^\alpha}\biggr|_\gamma = -u_{\bar{\alpha}}, \qquad \frac{\partial \hat{x}^a}{\partial x^\alpha}\biggr|_\gamma = e^a_{\bar{\alpha}}, \tag{9.10}$$

because in the limit $x \to \bar{x}$, the dual tetrad $(\bar{e}^0_\alpha, \bar{e}^a_\alpha)$ at x coincides with the dual tetrad $(-u_{\bar{\alpha}}, e^a_{\bar{\alpha}})$ at $\bar{x}$. It follows that on γ, the transformation matrix between the original coordinates x^α and the FNC $(t, \hat{x}^a)$ is formed by the Fermi-Walker transported tetrad:

$$\frac{\partial x^\alpha}{\partial t}\biggr|_\gamma = u^{\bar{\alpha}}, \qquad \frac{\partial x^\alpha}{\partial \hat{x}^a}\biggr|_\gamma = e^{\bar{\alpha}}_a. \tag{9.11}$$

This implies that the frame components of the acceleration vector, $a_a(t)$, are also the *components* of the acceleration vector in FNC; orthogonality between $u^{\bar{\alpha}}$ and $a^{\bar{\alpha}}$ means that $a_0 = 0$. Similarly, $R_{0c0d}(t)$, $R_{0cbd}(t)$, and $R_{acbd}(t)$ are the *components* of the Riemann tensor (evaluated on γ) in Fermi normal coordinates.

9.5 Metric near γ

Inversion of Eqs. (9.8) and (9.9) gives

$$\bar{e}^0_\alpha dx^\alpha = \biggl[1 + a_a\hat{x}^a + \frac{1}{2}R_{0c0d}\hat{x}^c\hat{x}^d + O(s^3)\biggr]\,dt + \biggl[\frac{1}{6}R_{0cbd}\hat{x}^c\hat{x}^d + O(s^3)\biggr]\,d\hat{x}^b \tag{9.12}$$

and

$$\bar{e}^a_\alpha dx^\alpha = \Big[\delta^a_{\ b} - \frac{1}{6}R^a_{\ cbd}\hat{x}^c\hat{x}^d + O(s^3)\Big]\, d\hat{x}^b + \Big[-\frac{1}{2}R^a_{\ c0d}\hat{x}^c\hat{x}^d + O(s^3)\Big]\, dt. \tag{9.13}$$

These relations, when specialized to the FNC, give the components of the dual tetrad at x. They can also be used to compute the metric at x, after invoking the completeness relations $g_{\alpha\beta} = -\bar{e}^0_\alpha\bar{e}^0_\beta + \delta_{ab}\bar{e}^a_\alpha\bar{e}^b_\beta$. This gives

$$g_{tt} = -\Big[1 + 2a_a\hat{x}^a + \big(a_a\hat{x}^a\big)^2 + R_{0c0d}\hat{x}^c\hat{x}^d + O(s^3)\Big], \tag{9.14}$$

$$g_{ta} = -\frac{2}{3}R_{0cad}\hat{x}^c\hat{x}^d + O(s^3), \tag{9.15}$$

$$g_{ab} = \delta_{ab} - \frac{1}{3}R_{acbd}\hat{x}^c\hat{x}^d + O(s^3). \tag{9.16}$$

This is the metric near γ in the Fermi normal coordinates. Recall that $a_a(t)$ are the components of the acceleration vector of γ — the timelike curve described by $\hat{x}^a = 0$ — while $R_{0c0d}(t)$, $R_{0cbd}(t)$, and $R_{acbd}(t)$ are the components of the Riemann tensor evaluated on γ.

Notice that on γ, the metric of Eqs. (9.14)–(9.16) reduces to $g_{tt} = -1$ and $g_{ab} = \delta_{ab}$. On the other hand, the nonvanishing Christoffel symbols (on γ) are $\Gamma^t_{\ ta} = \Gamma^a_{\ tt} = a_a$; these are zero (and the FNC enforce local flatness on the entire curve) when γ is a geodesic.

9.6 Thorne-Hartle-Zhang coordinates

The form of the metric can be simplified when the Ricci tensor vanishes on the world line:

$$R_{\mu\nu}(z) = 0. \tag{9.17}$$

In such circumstances, a transformation from the Fermi normal coordinates $(t, \hat{x}^a)$ to the *Thorne-Hartle-Zhang (THZ) coordinates* $(t, \hat{y}^a)$ brings the metric to the form

$$g_{tt} = -\Big[1 + 2a_a\hat{y}^a + \big(a_a\hat{y}^a\big)^2 + R_{0c0d}\hat{y}^c\hat{y}^d + O(s^3)\Big], \tag{9.18}$$

$$g_{ta} = -\frac{2}{3}R_{0cad}\hat{y}^c\hat{y}^d + O(s^3), \tag{9.19}$$

$$g_{ab} = \delta_{ab}\big(1 - R_{0c0d}\hat{y}^c\hat{y}^d\big) + O(s^3). \tag{9.20}$$

We see that the transformation leaves g_{tt} and g_{ta} unchanged, but that it diagonalizes g_{ab}. This metric was first displayed in Ref. [138] and the coordinate transformation was later produced by Zhang [139].

The key to the simplification comes from Eq. (9.17), which dramatically reduces the number of independent components of the Riemann tensor. In particular, Eq. (9.17) implies that the frame components R_{acbd} of the Riemann tensor are completely determined by $\mathcal{E}_{ab} := R_{0a0b}$, which in this special case is a symmetric-tracefree tensor. To prove this we invoke the completeness relations of Eq. (9.4) and take frame components of Eq. (9.17). This produces the three independent equations

$$\delta^{cd}R_{acbd} = \mathcal{E}_{ab}, \qquad \delta^{cd}R_{0cad} = 0, \qquad \delta^{cd}\mathcal{E}_{cd} = 0,$$

the last of which stating that $\mathcal{E}_{ab}$ has a vanishing trace. Taking the trace of the first equation gives $\delta^{ab}\delta^{cd}R_{acbd} = 0$, and this implies that R_{acbd} has five independent components. Since this is also the number of independent components of $\mathcal{E}_{ab}$, we see that the first equation can be inverted — R_{acbd} can be expressed in terms of $\mathcal{E}_{ab}$. A complete listing of the relevant relations is $R_{1212} = \mathcal{E}_{11} + \mathcal{E}_{22} = -\mathcal{E}_{33}$, $R_{1213} = \mathcal{E}_{23}$, $R_{1223} = -\mathcal{E}_{13}$, $R_{1313} = \mathcal{E}_{11} + \mathcal{E}_{33} = -\mathcal{E}_{22}$, $R_{1323} = \mathcal{E}_{12}$, and $R_{2323} = \mathcal{E}_{22} + \mathcal{E}_{33} = -\mathcal{E}_{11}$. These are summarized by

$$R_{acbd} = \delta_{ab}\mathcal{E}_{cd} + \delta_{cd}\mathcal{E}_{ab} - \delta_{ad}\mathcal{E}_{bc} - \delta_{bc}\mathcal{E}_{ad}, \tag{9.21}$$

and $\mathcal{E}_{ab} := R_{0a0b}$ satisfies $\delta^{ab}\mathcal{E}_{ab} = 0$.

We may also note that the relation $\delta^{cd}R_{0cad} = 0$, together with the usual symmetries of the Riemann tensor, imply that R_{0cad} too possesses five independent components. These may thus be related to another

symmetric-tracefree tensor $\mathcal{B}_{ab}$. We take the independent components to be $R_{0112} := -\mathcal{B}_{13}$, $R_{0113} := \mathcal{B}_{12}$, $R_{0123} := -\mathcal{B}_{11}$, $R_{0212} := -\mathcal{B}_{23}$, and $R_{0213} := \mathcal{B}_{22}$, and it is easy to see that all other components can be expressed in terms of these. For example, $R_{0223} = -R_{0113} = -\mathcal{B}_{12}$, $R_{0312} = -R_{0123} + R_{0213} = \mathcal{B}_{11} + \mathcal{B}_{22} = -\mathcal{B}_{33}$, $R_{0313} = -R_{0212} = \mathcal{B}_{23}$, and $R_{0323} = R_{0112} = -\mathcal{B}_{13}$. These relations are summarized by

$$R_{0abc} = -\varepsilon_{bcd}\mathcal{B}^{d}{}_{a}, \tag{9.22}$$

where ε_{abc} is the three-dimensional permutation symbol. The inverse relation is $\mathcal{B}^{a}{}_{b} = -\frac{1}{2}\varepsilon^{acd}R_{0bcd}$.

Substitution of Eq. (9.21) into Eq. (9.16) gives

$$g_{ab} = \delta_{ab}\Bigl(1 - \frac{1}{3}\mathcal{E}_{cd}\hat{x}^c\hat{x}^d\Bigr) - \frac{1}{3}\bigl(\hat{x}_c\hat{x}^c\bigr)\mathcal{E}_{ab} + \frac{1}{3}\hat{x}_a\mathcal{E}_{bc}\hat{x}^c + \frac{1}{3}\hat{x}_b\mathcal{E}_{ac}\hat{x}^c + O(s^3),$$

and we have not yet achieved the simple form of Eq. (9.20). The missing step is the transformation from the FNC $\hat{x}^a$ to the THZ coordinates $\hat{y}^a$. This is given by

$$\hat{y}^a = \hat{x}^a + \xi^a, \qquad \xi^a = -\frac{1}{6}\bigl(\hat{x}_c\hat{x}^c\bigr)\mathcal{E}_{ab}\hat{x}^b + \frac{1}{3}\hat{x}_a\mathcal{E}_{bc}\hat{x}^b\hat{x}^c + O(s^4). \tag{9.23}$$

It is easy to see that this transformation does not affect g_{tt} nor g_{ta} at orders s and s^2. The remaining components of the metric, however, transform according to $g_{ab}(\mathrm{THZ}) = g_{ab}(\mathrm{FNC}) - \xi_{a;b} - \xi_{b;a}$, where

$$\xi_{a;b} = \frac{1}{3}\delta_{ab}\mathcal{E}_{cd}\hat{x}^c\hat{x}^d - \frac{1}{6}\bigl(\hat{x}_c\hat{x}^c\bigr)\mathcal{E}_{ab} - \frac{1}{3}\mathcal{E}_{ac}\hat{x}^c\hat{x}_b + \frac{2}{3}\hat{x}_a\mathcal{E}_{bc}\hat{x}^c + O(s^3).$$

It follows that $g^{\mathrm{THZ}}_{ab} = \delta_{ab}(1 - \mathcal{E}_{cd}\hat{y}^c\hat{y}^d) + O(\hat{y}^3)$, which is just the same statement as in Eq. (9.20).

Alternative expressions for the components of the THZ metric are

$$g_{tt} = -\Bigl[1 + 2a_a\hat{y}^a + \bigl(a_a\hat{y}^a\bigr)^2 + \mathcal{E}_{ab}\hat{y}^a\hat{y}^b + O(s^3)\Bigr], \tag{9.24}$$

$$g_{ta} = -\frac{2}{3}\varepsilon_{abc}\mathcal{B}^{b}{}_{d}\hat{y}^c\hat{y}^d + O(s^3), \tag{9.25}$$

$$g_{ab} = \delta_{ab}\bigl(1 - \mathcal{E}_{cd}\hat{y}^c\hat{y}^d\bigr) + O(s^3). \tag{9.26}$$

10 Retarded coordinates

10.1 Geometrical elements

We introduce the same geometrical elements as in Sec. 9: we have a timelike curve γ described by relations $z^\mu(\tau)$, its normalized tangent vector $u^\mu = dz^\mu/d\tau$, and its acceleration vector $a^\mu = Du^\mu/d\tau$. We also have an orthonormal triad e^μ_a that is FW transported on the world line according to

$$\frac{De^\mu_a}{d\tau} = a_a u^\mu, \tag{10.1}$$

where $a_a(\tau) = a_\mu e^\mu_a$ are the frame components of the acceleration vector. Finally, we have a dual tetrad (e^0_μ, e^a_μ), with $e^0_\mu = -u_\mu$ and $e^a_\mu = \delta^{ab}g_{\mu\nu}e^\nu_b$. The tetrad and its dual give rise to the completeness relations

$$g^{\mu\nu} = -u^\mu u^\nu + \delta^{ab}e^\mu_a e^\nu_b, \qquad g_{\mu\nu} = -e^0_\mu e^0_\nu + \delta_{ab}\, e^a_\mu e^b_\nu, \tag{10.2}$$

which are the same as in Eq. (9.4).

The Fermi normal coordinates of Sec. 9 were constructed on the basis of a spacelike geodesic connecting a field point x to the world line. The retarded coordinates are based instead on a *null geodesic* going from the world line to the field point. We thus let x be within the normal convex neighbourhood of γ, β be the unique future-directed null geodesic that goes from the world line to x, and $x' := z(u)$ be the point at which β intersects the world line, with u denoting the value of the proper-time parameter at this point.

From the tetrad at x' we obtain another tetrad (e^α_0, e^α_a) at x by parallel transport on β. By raising the frame index and lowering the vectorial index we also obtain a dual tetrad at x: $e^0_\alpha = -g_{\alpha\beta} e^\beta_0$ and $e^a_\alpha = \delta^{ab} g_{\alpha\beta} e^\beta_b$. The metric at x can be then be expressed as

$$g_{\alpha\beta} = -e^0_\alpha e^0_\beta + \delta_{ab} e^a_\alpha e^b_\beta, \tag{10.3}$$

and the parallel propagator from x' to x is given by

$$g^\alpha_{\ \alpha'}(x,x') = -e^\alpha_0 u_{\alpha'} + e^\alpha_a e^a_{\alpha'}, \qquad g^{\alpha'}_{\ \alpha}(x',x) = u^{\alpha'} e^0_\alpha + e^{\alpha'}_a e^a_\alpha. \tag{10.4}$$

10.2 Definition of the retarded coordinates

The quasi-Cartesian version of the retarded coordinates are defined by

$$\hat{x}^0 = u, \qquad \hat{x}^a = -e^a_{\alpha'}(x')\sigma^{\alpha'}(x,x'), \qquad \sigma(x,x') = 0; \tag{10.5}$$

the last statement indicates that x' and x are linked by a null geodesic. From the fact that $\sigma^{\alpha'}$ is a null vector we obtain

$$r := (\delta_{ab}\hat{x}^a\hat{x}^b)^{1/2} = u_{\alpha'}\sigma^{\alpha'}, \tag{10.6}$$

and r is a positive quantity by virtue of the fact that β is a future-directed null geodesic — this makes $\sigma^{\alpha'}$ past-directed. In flat spacetime, $\sigma^{\alpha'} = -(x-x')^\alpha$, and in a Lorentz frame that is momentarily comoving with the world line, $r = t - t' > 0$; with the speed of light set equal to unity, r is also the spatial distance between x' and x as measured in this frame. In curved spacetime, the quantity $r = u_{\alpha'}\sigma^{\alpha'}$ can still be called the *retarded distance* between the point x and the world line. Another consequence of Eq. (10.5) is that

$$\sigma^{\alpha'} = -r\bigl(u^{\alpha'} + \Omega^a e^{\alpha'}_a\bigr), \tag{10.7}$$

where $\Omega^a := \hat{x}^a/r$ is a unit spatial vector that satisfies $\delta_{ab}\Omega^a\Omega^b = 1$.

A straightforward calculation reveals that under a displacement of the point x, the retarded coordinates change according to

$$du = -k_\alpha\,dx^\alpha, \qquad d\hat{x}^a = -\bigl(ra^a - \omega^a_{\ b}\hat{x}^b + e^a_{\alpha'}\sigma^{\alpha'}_{\ \beta'}u^{\beta'}\bigr)\,du - e^a_{\alpha'}\sigma^{\alpha'}_{\ \beta}\,dx^\beta, \tag{10.8}$$

where $k_\alpha = \sigma_\alpha/r$ is a future-directed null vector at x that is tangent to the geodesic β. To obtain these results we must keep in mind that a displacement of x typically induces a simultaneous displacement of x' because the new points $x+\delta x$ and $x'+\delta x'$ must also be linked by a null geodesic. We therefore have $0 = \sigma(x+\delta x, x'+\delta x') = \sigma_\alpha\,\delta x^\alpha + \sigma_{\alpha'}\,\delta x^{\alpha'}$, and the first relation of Eq. (10.8) follows from the fact that a displacement along the world line is described by $\delta x^{\alpha'} = u^{\alpha'}\,\delta u$.

10.3 The scalar field $r(x)$ and the vector field $k^\alpha(x)$

If we keep x' linked to x by the relation $\sigma(x,x') = 0$, then the quantity

$$r(x) = \sigma_{\alpha'}(x,x')u^{\alpha'}(x') \tag{10.9}$$

can be viewed as an ordinary scalar field defined in a neighbourhood of γ. We can compute the gradient of r by finding how r changes under a displacement of x (which again induces a displacement of x'). The result is

$$\partial_\beta r = -\bigl(\sigma_{\alpha'}a^{\alpha'} + \sigma_{\alpha'\beta'}u^{\alpha'}u^{\beta'}\bigr)k_\beta + \sigma_{\alpha'\beta}u^{\alpha'}. \tag{10.10}$$

Similarly, we can view

$$k^\alpha(x) = \frac{\sigma^\alpha(x,x')}{r(x)} \tag{10.11}$$

as an ordinary vector field, which is tangent to the congruence of null geodesics that emanate from x'. It is easy to check that this vector satisfies the identities

$$\sigma_{\alpha\beta}k^\beta = k_\alpha, \qquad \sigma_{\alpha'\beta}k^\beta = \frac{\sigma_{\alpha'}}{r}, \tag{10.12}$$

from which we also obtain $\sigma_{\alpha'\beta} u^{\alpha'} k^\beta = 1$. From this last result and Eq. (10.10) we deduce the important relation

$$k^\alpha \partial_\alpha r = 1. \tag{10.13}$$

In addition, combining the general statement $\sigma^\alpha = -g^\alpha_{\ \alpha'} \sigma^{\alpha'}$ — cf. Eq. (5.12) — with Eq. (10.7) gives

$$k^\alpha = g^\alpha_{\ \alpha'} \bigl(u^{\alpha'} + \Omega^a e^{\alpha'}_a\bigr); \tag{10.14}$$

the vector at x is therefore obtained by parallel transport of $u^{\alpha'} + \Omega^a e^{\alpha'}_a$ on β. From this and Eq. (10.4) we get the alternative expression

$$k^\alpha = e^\alpha_0 + \Omega^a e^\alpha_a, \tag{10.15}$$

which confirms that k^α is a future-directed null vector field (recall that $\Omega^a = \hat{x}^a/r$ is a unit vector).

The covariant derivative of k_α can be computed by finding how the vector changes under a displacement of x. (It is in fact easier to calculate first how rk_α changes, and then substitute our previous expression for $\partial_\beta r$.) The result is

$$rk_{\alpha;\beta} = \sigma_{\alpha\beta} - k_\alpha \sigma_{\beta\gamma'} u^{\gamma'} - k_\beta \sigma_{\alpha\gamma'} u^{\gamma'} + \bigl(\sigma_{\alpha'} a^{\alpha'} + \sigma_{\alpha'\beta'} u^{\alpha'} u^{\beta'}\bigr) k_\alpha k_\beta. \tag{10.16}$$

From this we infer that k^α satisfies the geodesic equation in affine-parameter form, $k^\alpha_{\ ;\beta} k^\beta = 0$, and Eq. (10.13) informs us that the affine parameter is in fact r. A displacement along a member of the congruence is therefore given by $dx^\alpha = k^\alpha\, dr$. Specializing to retarded coordinates, and using Eqs. (10.8) and (10.12), we find that this statement becomes $du = 0$ and $d\hat{x}^a = (\hat{x}^a/r)\, dr$, which integrate to $u = \text{constant}$ and $\hat{x}^a = r\Omega^a$, respectively, with Ω^a still denoting a constant unit vector. We have found that the congruence of null geodesics emanating from x' is described by

$$u = \text{constant}, \qquad \hat{x}^a = r\Omega^a(\theta^A) \tag{10.17}$$

in the retarded coordinates. Here, the two angles θ^A $(A = 1, 2)$ serve to parameterize the unit vector Ω^a, which is independent of r.

Equation (10.16) also implies that the expansion of the congruence is given by

$$\theta = k^\alpha_{\ ;\alpha} = \frac{\sigma^\alpha_{\ \alpha} - 2}{r}. \tag{10.18}$$

Using Eq. (6.10), we find that this becomes $r\theta = 2 - \frac{1}{3} R_{\alpha'\beta'} \sigma^{\alpha'} \sigma^{\beta'} + O(r^3)$, or

$$r\theta = 2 - \frac{1}{3} r^2 \bigl(R_{00} + 2R_{0a}\Omega^a + R_{ab}\Omega^a\Omega^b\bigr) + O(r^3) \tag{10.19}$$

after using Eq. (10.7). Here, $R_{00} = R_{\alpha'\beta'} u^{\alpha'} u^{\beta'}$, $R_{0a} = R_{\alpha'\beta'} u^{\alpha'} e^{\beta'}_a$, and $R_{ab} = R_{\alpha'\beta'} e^{\alpha'}_a e^{\beta'}_b$ are the frame components of the Ricci tensor evaluated at x'. This result confirms that the congruence is singular at $r = 0$, because θ diverges as $2/r$ in this limit; the caustic coincides with the point x'.

Finally, we infer from Eq. (10.16) that k^α is hypersurface orthogonal. This, together with the property that k^α satisfies the geodesic equation in affine-parameter form, implies that there exists a scalar field $u(x)$ such that

$$k_\alpha = -\partial_\alpha u. \tag{10.20}$$

This scalar field was already identified in Eq. (10.8): it is numerically equal to the proper-time parameter of the world line at x'. We conclude that the geodesics to which k^α is tangent are the generators of the null cone $u = \text{constant}$. As Eq. (10.17) indicates, a specific generator is selected by choosing a direction Ω^a (which can be parameterized by two angles θ^A), and r is an affine parameter on each generator. The geometrical meaning of the retarded coordinates is now completely clear; it is illustrated in Fig. 7.

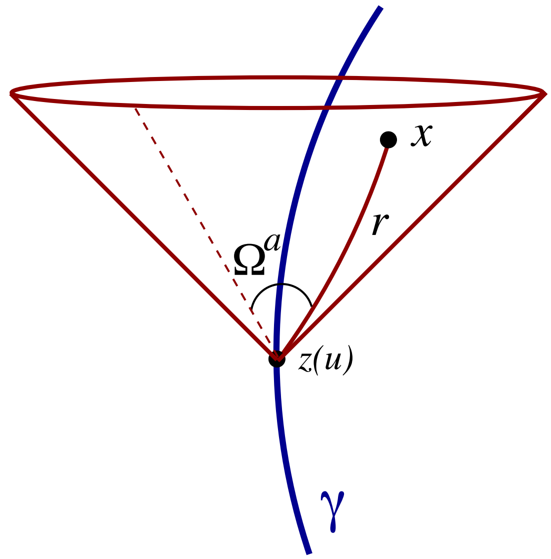

Figure 7: Retarded coordinates of a point x relative to a world line γ. The retarded time u selects a particular null cone, the unit vector $\Omega^a := \hat{x}^a/r$ selects a particular generator of this null cone, and the retarded distance r selects a particular point on this generator. This figure is identical to Fig. 4.

10.4 Frame components of tensor fields on the world line

The metric at x in the retarded coordinates will be expressed in terms of frame components of vectors and tensors evaluated on the world line γ. For example, if $a^{\alpha'}$ is the acceleration vector at x', then as we have seen,

$$a_a(u) = a_{\alpha'}\, e^{\alpha'}_a \tag{10.21}$$

are the frame components of the acceleration at proper time u.

Similarly,

$$R_{a0b0}(u) = R_{\alpha'\gamma'\beta'\delta'}\, e^{\alpha'}_a u^{\gamma'} e^{\beta'}_b u^{\delta'}, \quad R_{a0bd}(u) = R_{\alpha'\gamma'\beta'\delta'}\, e^{\alpha'}_a u^{\gamma'} e^{\beta'}_b e^{\delta'}_d, \quad R_{acbd}(u) = R_{\alpha'\gamma'\beta'\delta'}\, e^{\alpha'}_a e^{\gamma'}_c e^{\beta'}_b e^{\delta'}_d \tag{10.22}$$

are the frame components of the Riemann tensor evaluated on γ. From these we form the useful combinations

$$S_{ab}(u,\theta^A) = R_{a0b0} + R_{a0bc}\Omega^c + R_{b0ac}\Omega^c + R_{acbd}\Omega^c\Omega^d = S_{ba}, \tag{10.23}$$
$$S_a(u,\theta^A) = S_{ab}\Omega^b = R_{a0b0}\Omega^b - R_{ab0c}\Omega^b\Omega^c, \tag{10.24}$$
$$S(u,\theta^A) = S_a\Omega^a = R_{a0b0}\Omega^a\Omega^b, \tag{10.25}$$

in which the quantities $\Omega^a := \hat{x}^a/r$ depend on the angles θ^A only — they are independent of u and r.

We have previously introduced the frame components of the Ricci tensor in Eq. (10.19). The identity

$$R_{00} + 2R_{0a}\Omega^a + R_{ab}\Omega^a\Omega^b = \delta^{ab}S_{ab} - S \tag{10.26}$$

follows easily from Eqs. (10.23)–(10.25) and the definition of the Ricci tensor.

In Sec. 9 we saw that the frame components of a given tensor were also the components of this tensor (evaluated on the world line) in the Fermi normal coordinates. We should not expect this property to be true also in the case of the retarded coordinates: *the frame components of a tensor are not to be identified with the components of this tensor in the retarded coordinates.* The reason is that the retarded coordinates are in fact *singular* on the world line. As we shall see, they give rise to a metric that possesses a directional ambiguity at $r = 0$. (This can easily be seen in Minkowski spacetime by performing the coordinate transformation $u = t - \sqrt{x^2 + y^2 + z^2}$.) Components of tensors are therefore not defined on the world line, although they are perfectly well defined for $r \neq 0$. Frame components, on the other hand, are well defined both off and on the world line, and working with them will eliminate any difficulty associated with the singular nature of the retarded coordinates.

10.5 Coordinate displacements near γ

The changes in the quasi-Cartesian retarded coordinates under a displacement of x are given by Eq. (10.8). In these we substitute the standard expansions for $\sigma_{\alpha'\beta'}$ and $\sigma_{\alpha'\beta}$, as given by Eqs. (6.7) and (6.8), as well as Eqs. (10.7) and (10.14). After a straightforward (but fairly lengthy) calculation, we obtain the following expressions for the coordinate displacements:

$$du = \bigl(e^0_\alpha\, dx^\alpha\bigr) - \Omega_a \bigl(e^b_\alpha\, dx^\alpha\bigr), \tag{10.27}$$

$$\begin{aligned} d\hat{x}^a &= -\Bigl[r a^a + \frac{1}{2} r^2 S^a + O(r^3)\Bigr]\bigl(e^0_\alpha\, dx^\alpha\bigr) \\ &\quad + \Bigl[\delta^a_{\ b} + \Bigl(r a^a + \frac{1}{3} r^2 S^a\Bigr)\Omega_b + \frac{1}{6} r^2 S^a_{\ b} + O(r^3)\Bigr]\bigl(e^b_\alpha\, dx^\alpha\bigr). \end{aligned} \tag{10.28}$$

Notice that the result for du is exact, but that $d\hat{x}^a$ is expressed as an expansion in powers of r.

These results can also be expressed in the form of gradients of the retarded coordinates:

$$\partial_\alpha u = e^0_\alpha - \Omega_a e^a_\alpha, \tag{10.29}$$

$$\begin{aligned} \partial_\alpha \hat{x}^a &= -\Bigl[r a^a + \frac{1}{2} r^2 S^a + O(r^3)\Bigr] e^0_\alpha \\ &\quad + \Bigl[\delta^a_{\ b} + \Bigl(r a^a + \frac{1}{3} r^2 S^a\Bigr)\Omega_b + \frac{1}{6} r^2 S^a_{\ b} + O(r^3)\Bigr] e^b_\alpha. \end{aligned} \tag{10.30}$$

Notice that Eq. (10.29) follows immediately from Eqs. (10.15) and (10.20). From Eq. (10.30) and the identity $\partial_\alpha r = \Omega_a \partial_\alpha \hat{x}^a$ we also infer

$$\partial_\alpha r = -\Bigl[r a_a \Omega^a + \frac{1}{2} r^2 S + O(r^3)\Bigr] e^0_\alpha + \Bigl[\Bigl(1 + r a_b \Omega^b + \frac{1}{3} r^2 S\Bigr)\Omega_a + \frac{1}{6} r^2 S_a + O(r^3)\Bigr] e^a_\alpha, \tag{10.31}$$

where we have used the facts that $S_a = S_{ab}\Omega^b$ and $S = S_a \Omega^a$; these last results were derived in Eqs. (10.24) and (10.25). It may be checked that Eq. (10.31) agrees with Eq. (10.10).

10.6 Metric near γ

It is straightforward (but fairly tedious) to invert the relations of Eqs. (10.27) and (10.28) and solve for $e^0_\alpha\, dx^\alpha$ and $e^a_\alpha\, dx^\alpha$. The results are

$$e^0_\alpha\, dx^\alpha = \Bigl[1 + r a_a \Omega^a + \frac{1}{2} r^2 S + O(r^3)\Bigr] du + \Bigl[\Bigl(1 + \frac{1}{6} r^2 S\Bigr)\Omega_a - \frac{1}{6} r^2 S_a + O(r^3)\Bigr] d\hat{x}^a, \tag{10.32}$$

$$e^a_\alpha\, dx^\alpha = \Bigl[r a^a + \frac{1}{2} r^2 S^a + O(r^3)\Bigr] du + \Bigl[\delta^a_{\ b} - \frac{1}{6} r^2 S^a_{\ b} + \frac{1}{6} r^2 S^a \Omega_b + O(r^3)\Bigr] d\hat{x}^b. \tag{10.33}$$

These relations, when specialized to the retarded coordinates, give us the components of the dual tetrad (e^0_α, e^a_α) at x. The metric is then computed by using the completeness relations of Eq. (10.3). We find

$$g_{uu} = -\bigl(1 + r a_a \Omega^a\bigr)^2 + r^2 a^2 - r^2 S + O(r^3), \tag{10.34}$$

$$g_{ua} = -\Bigl(1 + r a_b \Omega^b + \frac{2}{3} r^2 S\Bigr)\Omega_a + r a_a + \frac{2}{3} r^2 S_a + O(r^3), \tag{10.35}$$

$$g_{ab} = \delta_{ab} - \Bigl(1 + \frac{1}{3} r^2 S\Bigr)\Omega_a \Omega_b - \frac{1}{3} r^2 S_{ab} + \frac{1}{3} r^2 \bigl(S_a \Omega_b + \Omega_a S_b\bigr) + O(r^3), \tag{10.36}$$

where $a^2 := \delta_{ab} a^a a^b$. We see (as was pointed out in Sec. 10.4) that the metric possesses a directional ambiguity on the world line: the metric at $r = 0$ still depends on the vector $\Omega^a = \hat{x}^a / r$ that specifies the direction to the point x. The retarded coordinates are therefore singular on the world line, and tensor components cannot be defined on γ.

By setting $S_{ab} = S_a = S = 0$ in Eqs. (10.34)–(10.36) we obtain the metric of flat spacetime in the retarded coordinates. This we express as

$$\begin{aligned} \eta_{uu} &= -\bigl(1 + r a_a \Omega^a\bigr)^2 + r^2 a^2, \\ \eta_{ua} &= -\bigl(1 + r a_b \Omega^b\bigr)\Omega_a + r a_a, \\ \eta_{ab} &= \delta_{ab} - \Omega_a \Omega_b. \end{aligned} \tag{10.37}$$

In spite of the directional ambiguity, the metric of flat spacetime has a unit determinant everywhere, and it is easily inverted:

$$\eta^{uu} = 0, \qquad \eta^{ua} = -\Omega^a, \qquad \eta^{ab} = \delta^{ab} + r\bigl(a^a\Omega^b + \Omega^a a^b\bigr). \tag{10.38}$$

The inverse metric also is ambiguous on the world line.

To invert the curved-spacetime metric of Eqs. (10.34)–(10.36) we express it as $g_{\alpha\beta} = \eta_{\alpha\beta} + h_{\alpha\beta} + O(r^3)$ and treat $h_{\alpha\beta} = O(r^2)$ as a perturbation. The inverse metric is then $g^{\alpha\beta} = \eta^{\alpha\beta} - \eta^{\alpha\gamma}\eta^{\beta\delta}h_{\gamma\delta} + O(r^3)$, or

$$g^{uu} = 0, \tag{10.39}$$

$$g^{ua} = -\Omega^a, \tag{10.40}$$

$$g^{ab} = \delta^{ab} + r\bigl(a^a\Omega^b + \Omega^a a^b\bigr) + \frac{1}{3}r^2 S^{ab} + \frac{1}{3}r^2\bigl(S^a\Omega^b + \Omega^a S^b\bigr) + O(r^3). \tag{10.41}$$

The results for g^{uu} and g^{ua} are exact, and they follow from the general relations $g^{\alpha\beta}(\partial_\alpha u)(\partial_\beta u) = 0$ and $g^{\alpha\beta}(\partial_\alpha u)(\partial_\beta r) = -1$ that are derived from Eqs. (10.13) and (10.20).

The metric determinant is computed from $\sqrt{-g} = 1 + \frac{1}{2}\eta^{\alpha\beta}h_{\alpha\beta} + O(r^3)$, which gives

$$\sqrt{-g} = 1 - \frac{1}{6}r^2\bigl(\delta^{ab}S_{ab} - S\bigr) + O(r^3) = 1 - \frac{1}{6}r^2\bigl(R_{00} + 2R_{0a}\Omega^a + R_{ab}\Omega^a\Omega^b\bigr) + O(r^3), \tag{10.42}$$

where we have substituted the identity of Eq. (10.26). Comparison with Eq. (10.19) gives us the interesting relation $\sqrt{-g} = \frac{1}{2}r\theta + O(r^3)$, where θ is the expansion of the generators of the null cones $u =$ constant.

10.7 Transformation to angular coordinates

Because the vector $\Omega^a = \hat{x}^a/r$ satisfies $\delta_{ab}\Omega^a\Omega^b = 1$, it can be parameterized by two angles θ^A. A canonical choice for the parameterization is $\Omega^a = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. It is then convenient to perform a coordinate transformation from $\hat{x}^a$ to (r, θ^A), using the relations $\hat{x}^a = r\Omega^a(\theta^A)$. (Recall from Sec. 10.3 that the angles θ^A are constant on the generators of the null cones $u =$ constant, and that r is an affine parameter on these generators. The relations $\hat{x}^a = r\Omega^a$ therefore describe the behaviour of the generators.) The differential form of the coordinate transformation is

$$d\hat{x}^a = \Omega^a\, dr + r\Omega^a_A\, d\theta^A, \tag{10.43}$$

where the transformation matrix

$$\Omega^a_A := \frac{\partial\Omega^a}{\partial\theta^A} \tag{10.44}$$

satisfies the identity $\Omega_a\Omega^a_A = 0$.

We introduce the quantities

$$\Omega_{AB} := \delta_{ab}\Omega^a_A\Omega^b_B, \tag{10.45}$$

which act as a (nonphysical) metric in the subspace spanned by the angular coordinates. In the canonical parameterization, $\Omega_{AB} = \mathrm{diag}(1, \sin^2\theta)$. We use the inverse of Ω_{AB}, denoted Ω^{AB}, to raise upper-case latin indices. We then define the new object

$$\Omega^A_a := \delta_{ab}\Omega^{AB}\Omega^b_B \tag{10.46}$$

which satisfies the identities

$$\Omega^A_a\Omega^a_B = \delta^A_B, \qquad \Omega^a_A\Omega^A_b = \delta^a{}_b - \Omega^a\Omega_b. \tag{10.47}$$

The second result follows from the fact that both sides are simultaneously symmetric in a and b, orthogonal to Ω_a and Ω^b, and have the same trace.

From the preceding results we establish that the transformation from $\hat{x}^a$ to (r, θ^A) is accomplished by

$$\frac{\partial\hat{x}^a}{\partial r} = \Omega^a, \qquad \frac{\partial\hat{x}^a}{\partial\theta^A} = r\Omega^a_A, \tag{10.48}$$

while the transformation from (r, θ^A) to $\hat{x}^a$ is accomplished by

$$\frac{\partial r}{\partial\hat{x}^a} = \Omega_a, \qquad \frac{\partial\theta^A}{\partial\hat{x}^a} = \frac{1}{r}\Omega^A_a. \tag{10.49}$$

With these transformation rules it is easy to show that in the angular coordinates, the metric is given by

$$g_{uu} = -\bigl(1 + r a_a \Omega^a\bigr)^2 + r^2 a^2 - r^2 S + O(r^3), \tag{10.50}$$

$$g_{ur} = -1, \tag{10.51}$$

$$g_{uA} = r\Bigl[r a_a + \frac{2}{3} r^2 S_a + O(r^3)\Bigr]\Omega^a_A, \tag{10.52}$$

$$g_{AB} = r^2\Bigl[\Omega_{AB} - \frac{1}{3} r^2 S_{ab}\Omega^a_A \Omega^b_B + O(r^3)\Bigr]. \tag{10.53}$$

The results $g_{ru} = -1$, $g_{rr} = 0$, and $g_{rA} = 0$ are exact, and they follow from the fact that in the retarded coordinates, $k_\alpha\, dx^\alpha = -du$ and $k^\alpha \partial_\alpha = \partial_r$.

The nonvanishing components of the inverse metric are

$$g^{ur} = -1, \tag{10.54}$$

$$g^{rr} = 1 + 2 r a_a \Omega^a + r^2 S + O(r^3), \tag{10.55}$$

$$g^{rA} = \frac{1}{r}\Bigl[r a^a + \frac{2}{3} r^2 S^a + O(r^3)\Bigr]\Omega^A_a, \tag{10.56}$$

$$g^{AB} = \frac{1}{r^2}\Bigl[\Omega^{AB} + \frac{1}{3} r^2 S^{ab}\Omega^A_a \Omega^B_b + O(r^3)\Bigr]. \tag{10.57}$$

The results $g^{uu} = 0$, $g^{ur} = -1$, and $g^{uA} = 0$ are exact, and they follow from the same reasoning as before.

Finally, we note that in the angular coordinates, the metric determinant is given by

$$\sqrt{-g} = r^2\sqrt{\Omega}\Bigl[1 - \frac{1}{6} r^2\bigl(R_{00} + 2R_{0a}\Omega^a + R_{ab}\Omega^a\Omega^b\bigr) + O(r^3)\Bigr], \tag{10.58}$$

where Ω is the determinant of Ω_{AB}; in the canonical parameterization, $\sqrt{\Omega} = \sin\theta$.

10.8 Specialization to $a^\mu = 0 = R_{\mu\nu}$

In this subsection we specialize our previous results to a situation where γ is a geodesic on which the Ricci tensor vanishes. We therefore set $a^\mu = 0 = R_{\mu\nu}$ everywhere on γ.

We have seen in Sec. 9.6 that when the Ricci tensor vanishes on γ, all frame components of the Riemann tensor can be expressed in terms of the symmetric-tracefree tensors $\mathcal{E}_{ab}(u)$ and $\mathcal{B}_{ab}(u)$. The relations are $R_{a0b0} = \mathcal{E}_{ab}$, $R_{a0bc} = \varepsilon_{bcd}\mathcal{B}^d{}_a$, and $R_{acbd} = \delta_{ab}\mathcal{E}_{cd} + \delta_{cd}\mathcal{E}_{ab} - \delta_{ad}\mathcal{E}_{bc} - \delta_{bc}\mathcal{E}_{ad}$. These can be substituted into Eqs. (10.23)–(10.25) to give

$$S_{ab}(u,\theta^A) = 2\mathcal{E}_{ab} - \Omega_a\mathcal{E}_{bc}\Omega^c - \Omega_b\mathcal{E}_{ac}\Omega^c + \delta_{ab}\mathcal{E}_{bc}\Omega^c\Omega^d + \varepsilon_{acd}\Omega^c\mathcal{B}^d{}_b + \varepsilon_{bcd}\Omega^c\mathcal{B}^d{}_a, \tag{10.59}$$

$$S_a(u,\theta^A) = \mathcal{E}_{ab}\Omega^b + \varepsilon_{abc}\Omega^b\mathcal{B}^c{}_d\Omega^d, \tag{10.60}$$

$$S(u,\theta^A) = \mathcal{E}_{ab}\Omega^a\Omega^b. \tag{10.61}$$

In these expressions the dependence on retarded time u is contained in $\mathcal{E}_{ab}$ and $\mathcal{B}_{ab}$, while the angular dependence is encoded in the unit vector Ω^a.

It is convenient to introduce the irreducible quantities

$$\mathcal{E}^* := \mathcal{E}_{ab}\Omega^a\Omega^b, \tag{10.62}$$

$$\mathcal{E}^*_a := \bigl(\delta_a{}^b - \Omega_a\Omega^b\bigr)\mathcal{E}_{bc}\Omega^c, \tag{10.63}$$

$$\mathcal{E}^*_{ab} := 2\mathcal{E}_{ab} - 2\Omega_a\mathcal{E}_{bc}\Omega^c - 2\Omega_b\mathcal{E}_{ac}\Omega^c + (\delta_{ab} + \Omega_a\Omega_b)\mathcal{E}^*, \tag{10.64}$$

$$\mathcal{B}^*_a := \varepsilon_{abc}\Omega^b\mathcal{B}^c{}_d\Omega^d, \tag{10.65}$$

$$\mathcal{B}^*_{ab} := \varepsilon_{acd}\Omega^c\mathcal{B}^d{}_e\bigl(\delta^e{}_b - \Omega^e\Omega_b\bigr) + \varepsilon_{bcd}\Omega^c\mathcal{B}^d{}_e\bigl(\delta^e{}_a - \Omega^e\Omega_a\bigr). \tag{10.66}$$

These are all orthogonal to Ω^a: $\mathcal{E}^*_a\Omega^a = \mathcal{B}^*_a\Omega^a = 0$ and $\mathcal{E}^*_{ab}\Omega^b = \mathcal{B}^*_{ab}\Omega^b = 0$. In terms of these Eqs. (10.59)–(10.61) become

$$S_{ab} = \mathcal{E}^*_{ab} + \Omega_a\mathcal{E}^*_b + \mathcal{E}^*_a\Omega_b + \Omega_a\Omega_b\mathcal{E}^* + \mathcal{B}^*_{ab} + \Omega_a\mathcal{B}^*_b + \mathcal{B}^*_a\Omega_b, \tag{10.67}$$

$$S_a = \mathcal{E}^*_a + \Omega_a\mathcal{E}^* + \mathcal{B}^*_a, \tag{10.68}$$

$$S = \mathcal{E}^*. \tag{10.69}$$

When Eqs. (10.67)–(10.69) are substituted into the metric tensor of Eqs. (10.34)–(10.36) — in which a_a is set equal to zero — we obtain the compact expressions

$$
\begin{aligned}
g_{uu} &= -1 - r^2\mathcal{E}^* + O(r^3), && (10.70)\\
g_{ua} &= -\Omega_a + \frac{2}{3}r^2\bigl(\mathcal{E}^*_a + \mathcal{B}^*_a\bigr) + O(r^3), && (10.71)\\
g_{ab} &= \delta_{ab} - \Omega_a\Omega_b - \frac{1}{3}r^2\bigl(\mathcal{E}^*_{ab} + \mathcal{B}^*_{ab}\bigr) + O(r^3). && (10.72)
\end{aligned}
$$

The metric becomes

$$
\begin{aligned}
g_{uu} &= -1 - r^2\mathcal{E}^* + O(r^3), && (10.73)\\
g_{ur} &= -1, && (10.74)\\
g_{uA} &= \frac{2}{3}r^3\bigl(\mathcal{E}^*_A + \mathcal{B}^*_A\bigr) + O(r^4), && (10.75)\\
g_{AB} &= r^2\Omega_{AB} - \frac{1}{3}r^4\bigl(\mathcal{E}^*_{AB} + \mathcal{B}^*_{AB}\bigr) + O(r^5) && (10.76)
\end{aligned}
$$

after transforming to angular coordinates using the rules of Eq. (10.48). Here we have introduced the projections

$$
\begin{aligned}
\mathcal{E}^*_A &:= \mathcal{E}^*_a\Omega^a_A = \mathcal{E}_{ab}\Omega^a_A\Omega^b, && (10.77)\\
\mathcal{E}^*_{AB} &:= \mathcal{E}^*_{ab}\Omega^a_A\Omega^b_B = 2\mathcal{E}_{ab}\Omega^a_A\Omega^b_B + \mathcal{E}^*\Omega_{AB}, && (10.78)\\
\mathcal{B}^*_A &:= \mathcal{B}^*_a\Omega^a_A = \varepsilon_{abc}\Omega^a_A\Omega^b\mathcal{B}^c{}_d\Omega^d, && (10.79)\\
\mathcal{B}^*_{AB} &:= \mathcal{B}^*_{ab}\Omega^a_A\Omega^b_B = 2\varepsilon_{acd}\Omega^c\mathcal{B}^d{}_b\Omega^a_{(A}\Omega^b_{B)}. && (10.80)
\end{aligned}
$$

It may be noted that the inverse relations are $\mathcal{E}^*_a = \mathcal{E}^*_A\Omega^A_a$, $\mathcal{B}^*_a = \mathcal{B}^*_A\Omega^A_a$, $\mathcal{E}^*_{ab} = \mathcal{E}^*_{AB}\Omega^A_a\Omega^B_b$, and $\mathcal{B}^*_{ab} = \mathcal{B}^*_{AB}\Omega^A_a\Omega^B_b$, where Ω^A_a was introduced in Eq. (10.46).

11 Transformation between Fermi and retarded coordinates; advanced point

A point x in the normal convex neighbourhood of a world line γ can be assigned a set of Fermi normal coordinates (as in Sec. 9), or it can be assigned a set of retarded coordinates (Sec. 10). These coordinate systems can obviously be related to one another, and our first task in this section (which will occupy us in Secs. 11.1–11.3) will be to derive the transformation rules. We begin by refining our notation so as to eliminate any danger of ambiguity.

The Fermi normal coordinates of x refer to a point $\bar{x} := z(t)$ on γ that is related to x by a spacelike geodesic that intersects γ orthogonally; see Fig. 8. We refer to this point as x's *simultaneous point*, and to tensors at $\bar{x}$ we assign indices $\bar{\alpha}$, $\bar{\beta}$, etc. We let $(t, s\omega^a)$ be the Fermi normal coordinates of x, with t denoting the value of γ's proper-time parameter at $\bar{x}$, $s = \sqrt{2\sigma(x,\bar{x})}$ representing the proper distance from $\bar{x}$ to x along the spacelike geodesic, and ω^a denoting a unit vector ($\delta_{ab}\omega^a\omega^b = 1$) that determines the direction of the geodesic. The Fermi normal coordinates are defined by $s\omega^a = -e^a_{\bar{\alpha}}\sigma^{\bar{\alpha}}$ and $\sigma_{\bar{\alpha}}u^{\bar{\alpha}} = 0$. Finally, we denote by $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ the tetrad at x that is obtained by parallel transport of $(u^{\bar{\alpha}}, e^{\bar{\alpha}}_a)$ on the spacelike geodesic.

The retarded coordinates of x refer to a point $x' := z(u)$ on γ that is linked to x by a future-directed null geodesic; see Fig. 8. We refer to this point as x's *retarded point*, and to tensors at x' we assign indices α', β', etc. We let $(u, r\Omega^a)$ be the retarded coordinates of x, with u denoting the value of γ's proper-time parameter at x', $r = \sigma_{\alpha'}u^{\alpha'}$ representing the affine-parameter distance from x' to x along the null geodesic, and Ω^a denoting a unit vector ($\delta_{ab}\Omega^a\Omega^b = 1$) that determines the direction of the geodesic. The retarded coordinates are defined by $r\Omega^a = -e^a_{\alpha'}\sigma^{\alpha'}$ and $\sigma(x,x') = 0$. Finally, we denote by (e^α_0, e^α_a) the tetrad at x that is obtained by parallel transport of $(u^{\alpha'}, e^{\alpha'}_a)$ on the null geodesic.

The reader who does not wish to follow the details of this discussion can be informed that: (i) our results concerning the transformation from the retarded coordinates (u, r, Ω^a) to the Fermi normal coordinates

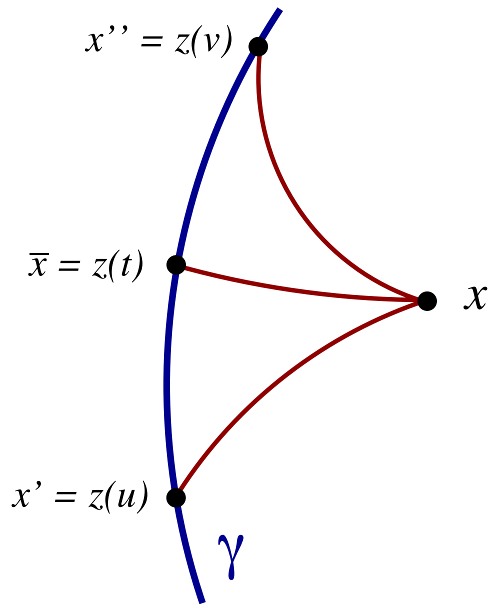

Figure 8: The retarded, simultaneous, and advanced points on a world line γ. The retarded point $x' := z(u)$ is linked to x by a future-directed null geodesic. The simultaneous point $\bar{x} := z(t)$ is linked to x by a spacelike geodesic that intersects γ orthogonally. The advanced point $x'' := z(v)$ is linked to x by a past-directed null geodesic.

(t, s, ω^a) are contained in Eqs. (11.1)–(11.3) below; (ii) our results concerning the transformation from the Fermi normal coordinates (t, s, ω^a) to the retarded coordinates (u, r, Ω^a) are contained in Eqs. (11.4)–(11.6); (iii) the decomposition of each member of $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ in the tetrad (e^α_0, e^α_a) is given in retarded coordinates by Eqs. (11.7) and (11.8); and (iv) the decomposition of each member of (e^α_0, e^α_a) in the tetrad $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ is given in Fermi normal coordinates by Eqs. (11.9) and (11.10).

Our final task will be to define, along with the retarded and simultaneous points, an *advanced point* x'' on the world line γ; see Fig. 8. This is taken on in Sec. 11.4.

11.1 From retarded to Fermi coordinates

Quantities at $\bar{x} := z(t)$ can be related to quantities at $x' := z(u)$ by Taylor expansion along the world line γ. To implement this strategy we must first find an expression for $\Delta := t - u$. (Although we use the same notation, this should not be confused with the van Vleck determinant introduced in Sec. 7.)

Consider the function $p(\tau)$ of the proper-time parameter τ defined by

$$p(\tau) = \sigma_\mu\big(x, z(\tau)\big)u^\mu(\tau),$$

in which x is kept fixed and in which $z(\tau)$ is an arbitrary point on the world line. We have that $p(u) = r$ and $p(t) = 0$, and Δ can ultimately be obtained by expressing $p(t)$ as $p(u + \Delta)$ and expanding in powers of Δ. Formally,

$$p(t) = p(u) + \dot{p}(u)\Delta + \frac{1}{2}\ddot{p}(u)\Delta^2 + \frac{1}{6}p^{(3)}(u)\Delta^3 + O(\Delta^4),$$

where overdots (or a number within brackets) indicate repeated differentiation with respect to τ. We have

$$\begin{aligned}
\dot{p}(u) &= \sigma_{\alpha'\beta'}u^{\alpha'}u^{\beta'} + \sigma_{\alpha'}a^{\alpha'},\\
\ddot{p}(u) &= \sigma_{\alpha'\beta'\gamma'}u^{\alpha'}u^{\beta'}u^{\gamma'} + 3\sigma_{\alpha'\beta'}u^{\alpha'}a^{\beta'} + \sigma_{\alpha'}\dot{a}^{\alpha'},\\
p^{(3)}(u) &= \sigma_{\alpha'\beta'\gamma'\delta'}u^{\alpha'}u^{\beta'}u^{\gamma'}u^{\delta'} + \sigma_{\alpha'\beta'\gamma'}\big(5a^{\alpha'}u^{\beta'}u^{\gamma'} + u^{\alpha'}u^{\beta'}a^{\gamma'}\big) + \sigma_{\alpha'\beta'}\big(3a^{\alpha'}a^{\beta'} + 4u^{\alpha'}\dot{a}^{\beta'}\big) + \sigma_{\alpha'}\ddot{a}^{\alpha'},
\end{aligned}$$

where $a^\mu = Du^\mu/d\tau$, $\dot{a}^\mu = Da^\mu/d\tau$, and $\ddot{a}^\mu = D\dot{a}^\mu/d\tau$.

We now express all of this in retarded coordinates by invoking the expansion of Eq. (6.7) for $\sigma_{\alpha'\beta'}$ (as well as additional expansions for the higher derivatives of the world function, obtained by further differentiation of this result) and the relation $\sigma^{\alpha'} = -r(u^{\alpha'} + \Omega^a e^{\alpha'}_a)$ first derived in Eq. (10.7). With a degree of accuracy sufficient for our purposes we obtain

$$\begin{aligned}
\dot{p}(u) &= -\Big[1 + ra_a\Omega^a + \frac{1}{3}r^2 S + O(r^3)\Big], \\
\ddot{p}(u) &= -r\big(\dot{a}_0 + \dot{a}_a\Omega^a\big) + O(r^2), \\
p^{(3)}(u) &= \dot{a}_0 + O(r),
\end{aligned}$$

where $S = R_{a0b0}\Omega^a\Omega^b$ was first introduced in Eq. (10.25), and where $\dot{a}_0 := \dot{a}_{\alpha'}u^{\alpha'}$, $\dot{a}_a := \dot{a}_{\alpha'}e^{\alpha'}_a$ are the frame components of the covariant derivative of the acceleration vector. To arrive at these results we made use of the identity $a_{\alpha'}a^{\alpha'} + \dot{a}_{\alpha'}u^{\alpha'} = 0$ that follows from the fact that a^μ is orthogonal to u^μ. Notice that there is no distinction between the two possible interpretations $\dot{a}_a := da_a/d\tau$ and $\dot{a}_a := \dot{a}_\mu e^\mu_a$ for the quantity $\dot{a}_a(\tau)$; their equality follows at once from the substitution of $De^\mu_a/d\tau = a_a u^\mu$ (which states that the basis vectors are Fermi-Walker transported on the world line) into the identity $da_a/d\tau = D(a_\nu e^\nu_a)/d\tau$.

Collecting our results we obtain

$$r = \Big[1 + ra_a\Omega^a + \frac{1}{3}r^2 S + O(r^3)\Big]\Delta + \frac{1}{2}r\Big[\dot{a}_0 + \dot{a}_a\Omega^a + O(r)\Big]\Delta^2 - \frac{1}{6}\Big[\dot{a}_0 + O(r)\Big]\Delta^3 + O(\Delta^4),$$

which can readily be solved for $\Delta := t - u$ expressed as an expansion in powers of r. The final result is

$$t = u + r\bigg\{1 - ra_a(u)\Omega^a + r^2\big[a_a(u)\Omega^a\big]^2 - \frac{1}{3}r^2\dot{a}_0(u) - \frac{1}{2}r^2\dot{a}_a(u)\Omega^a - \frac{1}{3}r^2 R_{a0b0}(u)\Omega^a\Omega^b + O(r^3)\bigg\}, \quad (11.1)$$

where we show explicitly that all frame components are evaluated at the retarded point $z(u)$.

To obtain relations between the spatial coordinates we consider the functions

$$p_a(\tau) = -\sigma_\mu\big(x, z(\tau)\big)e^\mu_a(\tau),$$

in which x is fixed and $z(\tau)$ is an arbitrary point on γ. We have that the retarded coordinates are given by $r\Omega^a = p^a(u)$, while the Fermi coordinates are given instead by $s\omega^a = p^a(t) = p^a(u+\Delta)$. This last expression can be expanded in powers of Δ, producing

$$s\omega^a = p^a(u) + \dot{p}^a(u)\Delta + \frac{1}{2}\ddot{p}^a(u)\Delta^2 + \frac{1}{6}p^{a(3)}(u)\Delta^3 + O(\Delta^4)$$

with

$$\begin{aligned}
\dot{p}_a(u) &= -\sigma_{\alpha'\beta'}e^{\alpha'}_a u^{\beta'} - \big(\sigma_{\alpha'}u^{\alpha'}\big)\big(a_{\beta'}e^{\beta'}_a\big) \\
&= -ra_a - \frac{1}{3}r^2 S_a + O(r^3), \\
\ddot{p}_a(u) &= -\sigma_{\alpha'\beta'\gamma'}e^{\alpha'}_a u^{\beta'}u^{\gamma'} - \big(2\sigma_{\alpha'\beta'}u^{\alpha'}u^{\beta'} + \sigma_{\alpha'}a^{\alpha'}\big)\big(a_{\gamma'}e^{\gamma'}_a\big) - \sigma_{\alpha'\beta'}e^{\alpha'}_a a^{\beta'} - \big(\sigma_{\alpha'}u^{\alpha'}\big)\big(\dot{a}_{\beta'}e^{\beta'}_a\big) \\
&= \big(1 + ra_b\Omega^b\big)a_a - r\dot{a}_a + \frac{1}{3}rR_{a0b0}\Omega^b + O(r^2), \\
p^{(3)}_a(u) &= -\sigma_{\alpha'\beta'\gamma'\delta'}e^{\alpha'}_a u^{\beta'}u^{\gamma'}u^{\delta'} - \big(3\sigma_{\alpha'\beta'\gamma'}u^{\alpha'}u^{\beta'}u^{\gamma'} + 6\sigma_{\alpha'\beta'}u^{\alpha'}a^{\beta'} + \sigma_{\alpha'}\dot{a}^{\alpha'} + \sigma_{\alpha'}u^{\alpha'}\dot{a}_{\beta'}u^{\beta'}\big)\big(a_{\delta'}e^{\delta'}_a\big) \\
&\quad - \sigma_{\alpha'\beta'\gamma'}e^{\alpha'}_a\big(2a^{\beta'}u^{\gamma'} + u^{\beta'}a^{\gamma'}\big) - \big(3\sigma_{\alpha'\beta'}u^{\alpha'}u^{\beta'} + 2\sigma_{\alpha'}a^{\alpha'}\big)\big(\dot{a}_{\gamma'}e^{\gamma'}_a\big) - \sigma_{\alpha'\beta'}e^{\alpha'}_a\dot{a}^{\beta'} \\
&\quad - \big(\sigma_{\alpha'}u^{\alpha'}\big)\big(\ddot{a}_{\beta'}e^{\beta'}_a\big) \\
&= 2\dot{a}_a + O(r).
\end{aligned}$$

To arrive at these results we have used the same expansions as before and re-introduced $S_a = R_{a0b0}\Omega^b - R_{ab0c}\Omega^b\Omega^c$, as it was first defined in Eq. (10.24).

Collecting our results we obtain

$$\begin{aligned}
s\omega^a &= r\Omega^a - r\Big[a^a + \frac{1}{3}rS^a + O(r^2)\Big]\Delta + \frac{1}{2}\Big[\big(1 + ra_b\Omega^b\big)a^a - r\dot{a}^a + \frac{1}{3}rR^a{}_{0b0}\Omega^b + O(r^2)\Big]\Delta^2 \\
&\quad + \frac{1}{3}\Big[\dot{a}^a + O(r)\Big]\Delta^3 + O(\Delta^4),
\end{aligned}$$

which becomes

$$s\omega^a = r\biggl\{\Omega^a - \frac{1}{2}r\bigl[1 - r a_b(u)\Omega^b\bigr]a^a(u) - \frac{1}{6}r^2\dot{a}^a(u) - \frac{1}{6}r^2 R^a_{\ 0b0}(u)\Omega^b + \frac{1}{3}r^2 R^a_{\ b0c}(u)\Omega^b\Omega^c + O(r^3)\biggr\} \quad (11.2)$$

after substituting Eq. (11.1) for $\Delta := t - u$. From squaring Eq. (11.2) and using the identity $\delta_{ab}\omega^a\omega^b = 1$ we can also deduce

$$s = r\biggl\{1 - \frac{1}{2}r a_a(u)\Omega^a + \frac{3}{8}r^2\bigl[a_a(u)\Omega^a\bigr]^2 - \frac{1}{8}r^2\dot{a}_0(u) - \frac{1}{6}r^2\dot{a}_a(u)\Omega^a - \frac{1}{6}r^2 R_{a0b0}(u)\Omega^a\Omega^b + O(r^3)\biggr\} \quad (11.3)$$

for the spatial distance between x and $z(t)$.

11.2 From Fermi to retarded coordinates

The techniques developed in the preceding subsection can easily be adapted to the task of relating the retarded coordinates of x to its Fermi normal coordinates. Here we use $\bar{x} := z(t)$ as the reference point and express all quantities at $x' := z(u)$ as Taylor expansions about $\tau = t$.

We begin by considering the function

$$\sigma(\tau) = \sigma\bigl(x, z(\tau)\bigr)$$

of the proper-time parameter τ on γ. We have that $\sigma(t) = \frac{1}{2}s^2$ and $\sigma(u) = 0$, and $\Delta := t - u$ is now obtained by expressing $\sigma(u)$ as $\sigma(t - \Delta)$ and expanding in powers of Δ. Using the fact that $\dot{\sigma}(\tau) = p(\tau)$, we have

$$\sigma(u) = \sigma(t) - p(t)\Delta + \frac{1}{2}\dot{p}(t)\Delta^2 - \frac{1}{6}\ddot{p}(t)\Delta^3 + \frac{1}{24}p^{(3)}(t)\Delta^4 + O(\Delta^5).$$

Expressions for the derivatives of $p(\tau)$ evaluated at $\tau = t$ can be constructed from results derived previously in Sec. 11.1: it suffices to replace all primed indices by barred indices and then substitute the relation $\sigma^{\bar{\alpha}} = -s\omega^a e^{\bar{\alpha}}_a$ that follows immediately from Eq. (9.5). This gives

$$\begin{aligned}
\dot{p}(t) &= -\Bigl[1 + s a_a\omega^a + \frac{1}{3}s^2 R_{a0b0}\omega^a\omega^b + O(s^3)\Bigr],\\
\ddot{p}(t) &= -s\dot{a}_a\omega^a + O(s^2),\\
p^{(3)}(t) &= \dot{a}_0 + O(s),
\end{aligned}$$

and then

$$s^2 = \Bigl[1 + s a_a\omega^a + \frac{1}{3}s^2 R_{a0b0}\omega^a\omega^b + O(s^3)\Bigr]\Delta^2 - \frac{1}{3}s\Bigl[\dot{a}_a\omega^a + O(s)\Bigr]\Delta^3 - \frac{1}{12}\Bigl[\dot{a}_0 + O(s)\Bigr]\Delta^4 + O(\Delta^5)$$

after recalling that $p(t) = 0$. Solving for Δ as an expansion in powers of s returns

$$u = t - s\biggl\{1 - \frac{1}{2}s a_a(t)\omega^a + \frac{3}{8}s^2\bigl[a_a(t)\omega^a\bigr]^2 + \frac{1}{24}s^2\dot{a}_0(t) + \frac{1}{6}s^2\dot{a}_a(t)\omega^a - \frac{1}{6}s^2 R_{a0b0}(t)\omega^a\omega^b + O(s^3)\biggr\}, \quad (11.4)$$

in which we emphasize that all frame components are evaluated at the simultaneous point $z(t)$.

An expression for $r = p(u)$ can be obtained by expanding $p(t - \Delta)$ in powers of Δ. We have

$$r = -\dot{p}(t)\Delta + \frac{1}{2}\ddot{p}(t)\Delta^2 - \frac{1}{6}p^{(3)}(t)\Delta^3 + O(\Delta^4),$$

and substitution of our previous results gives

$$r = s\biggl\{1 + \frac{1}{2}s a_a(t)\omega^a - \frac{1}{8}s^2\bigl[a_a(t)\omega^a\bigr]^2 - \frac{1}{8}s^2\dot{a}_0(t) - \frac{1}{3}s^2\dot{a}_a(t)\omega^a + \frac{1}{6}s^2 R_{a0b0}(t)\omega^a\omega^b + O(s^3)\biggr\} \quad (11.5)$$

for the retarded distance between x and $z(u)$.

Finally, the retarded coordinates $r\Omega^a = p^a(u)$ can be related to the Fermi coordinates by expanding $p^a(t-\Delta)$ in powers of Δ, so that

$$r\Omega^a = p^a(t) - \dot{p}^a(t)\Delta + \frac{1}{2}\ddot{p}^a(t)\Delta^2 - \frac{1}{6}p^{a(3)}(t)\Delta^3 + O(\Delta^4).$$

Results from the preceding subsection can again be imported with mild alterations, and we find

$$\begin{aligned}
\dot{p}_a(t) &= \frac{1}{3}s^2 R_{ab0c}\omega^b\omega^c + O(s^3), \\
\ddot{p}_a(t) &= \bigl(1 + s a_b\omega^b\bigr)a_a + \frac{1}{3}sR_{a0b0}\omega^b + O(s^2), \\
p_a^{(3)}(t) &= 2\dot{a}_a(t) + O(s).
\end{aligned}$$

This, together with Eq. (11.4), gives

$$r\Omega^a = s\biggl\{\omega^a + \frac{1}{2}sa^a(t) - \frac{1}{3}s^2\dot{a}^a(t) - \frac{1}{3}s^2R^a_{\ b0c}(t)\omega^b\omega^c + \frac{1}{6}s^2R^a_{\ 0b0}(t)\omega^b + O(s^3)\biggr\}. \qquad (11.6)$$

It may be checked that squaring this equation and using the identity $\delta_{ab}\Omega^a\Omega^b = 1$ returns the same result as Eq. (11.5).

11.3 Transformation of the tetrads at x

Recall that we have constructed two sets of basis vectors at x. The first set is the tetrad $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ that is obtained by parallel transport of $(u^{\bar{\alpha}}, e^{\bar{\alpha}}_a)$ on the spacelike geodesic that links x to the simultaneous point $\bar{x} := z(t)$. The second set is the tetrad (e^α_0, e^α_a) that is obtained by parallel transport of $(u^{\alpha'}, e^{\alpha'}_a)$ on the null geodesic that links x to the retarded point $x' := z(u)$. Since each tetrad forms a complete set of basis vectors, each member of $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ can be decomposed in the tetrad (e^α_0, e^α_a), and correspondingly, each member of (e^α_0, e^α_a) can be decomposed in the tetrad $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$. These decompositions are worked out in this subsection. For this purpose we shall consider the functions

$$p^\alpha(\tau) = g^\alpha_{\ \mu}\bigl(x, z(\tau)\bigr)u^\mu(\tau), \qquad p^\alpha_a(\tau) = g^\alpha_{\ \mu}\bigl(x, z(\tau)\bigr)e^\mu_a(\tau),$$

in which x is a fixed point in a neighbourhood of γ, $z(\tau)$ is an arbitrary point on the world line, and $g^\alpha_{\ \mu}(x,z)$ is the parallel propagator on the unique geodesic that links x to z. We have $\bar{e}^\alpha_0 = p^\alpha(t)$, $\bar{e}^\alpha_a = p^\alpha_a(t)$, $e^\alpha_0 = p^\alpha(u)$, and $e^\alpha_a = p^\alpha_a(u)$.

We begin with the decomposition of $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ in the tetrad (e^α_0, e^α_a) associated with the retarded point $z(u)$. This decomposition will be expressed in the retarded coordinates as an expansion in powers of r. As in Sec. 9.1 we express quantities at $z(t)$ in terms of quantities at $z(u)$ by expanding in powers of $\Delta := t - u$. We have

$$\bar{e}^\alpha_0 = p^\alpha(u) + \dot{p}^\alpha(u)\Delta + \frac{1}{2}\ddot{p}^\alpha(u)\Delta^2 + O(\Delta^3),$$

with

$$\begin{aligned}
\dot{p}^\alpha(u) &= g^\alpha_{\ \alpha';\beta'}u^{\alpha'}u^{\beta'} + g^\alpha_{\ \alpha'}a^{\alpha'} \\
&= \Bigl[a^a + \frac{1}{2}rR^a_{\ 0b0}\Omega^b + O(r^2)\Bigr]e^\alpha_a, \\
\ddot{p}^\alpha(u) &= g^\alpha_{\ \alpha';\beta'\gamma'}u^{\alpha'}u^{\beta'}u^{\gamma'} + g^\alpha_{\ \alpha';\beta'}\bigl(2a^{\alpha'}u^{\beta'} + u^{\alpha'}a^{\beta'}\bigr) + g^\alpha_{\ \alpha'}\dot{a}^{\alpha'} \\
&= \Bigl[-\dot{a}_0 + O(r)\Bigr]e^\alpha_0 + \Bigl[\dot{a}^a + O(r)\Bigr]e^\alpha_a,
\end{aligned}$$

where we have used the expansions of Eq. (6.11) as well as the decompositions of Eq. (10.4). Collecting these results and substituting Eq. (11.1) for Δ yields

$$\bar{e}^\alpha_0 = \Bigl[1 - \frac{1}{2}r^2\dot{a}_0(u) + O(r^3)\Bigr]e^\alpha_0 + \Bigl[r\bigl(1 - a_b\Omega^b\bigr)a^a(u) + \frac{1}{2}r^2\dot{a}^a(u) + \frac{1}{2}r^2R^a_{\ 0b0}(u)\Omega^b + O(r^3)\Bigr]e^\alpha_a. \qquad (11.7)$$

Similarly, we have

$$\bar{e}^\alpha_a = p^\alpha_a(u) + \dot{p}^\alpha_a(u)\Delta + \frac{1}{2}\ddot{p}^\alpha_a(u)\Delta^2 + O(\Delta^3),$$

with

$$\begin{aligned}
\dot{p}^\alpha_a(u) &= g^\alpha_{\ \alpha';\beta'} e^{\alpha'}_a u^{\beta'} + \bigl(g^\alpha_{\ \alpha'} u^{\alpha'}\bigr)\bigl(a_{\beta'} e^{\beta'}_a\bigr) \\
&= \Bigl[a_a + \frac{1}{2} r R_{a0b0}\Omega^b + O(r^2)\Bigr] e^\alpha_0 + \Bigl[-\frac{1}{2} r R^b_{\ a0c}\Omega^c + O(r^2)\Bigr] e^\alpha_b, \\
\ddot{p}^\alpha_a(u) &= g^\alpha_{\ \alpha';\beta'\gamma'} e^{\alpha'}_a u^{\beta'} u^{\gamma'} + g^\alpha_{\ \alpha';\beta'}\bigl(2u^{\alpha'} u^{\beta'} a_{\gamma'} e^{\gamma'}_a + e^{\alpha'}_a a^{\beta'}\bigr) + \bigl(g^\alpha_{\ \alpha'} a^{\alpha'}\bigr)\bigl(a_{\beta'} e^{\beta'}_a\bigr) + \bigl(g^\alpha_{\ \alpha'} u^{\alpha'}\bigr)\bigl(\dot{a}_{\beta'} e^{\beta'}_a\bigr) \\
&= \Bigl[\dot{a}_a + O(r)\Bigr] e^\alpha_0 + \Bigl[a_a a^b + O(r)\Bigr] e^\alpha_b,
\end{aligned}$$

and all this gives

$$\begin{aligned}
\bar{e}^\alpha_a &= \Bigl[\delta^b_{\ a} + \frac{1}{2} r^2 a^b(u) a_a(u) - \frac{1}{2} r^2 R^b_{\ a0c}(u)\Omega^c + O(r^3)\Bigr]\, e^\alpha_b \\
&\quad + \Bigl[r\bigl(1 - r a_b \Omega^b\bigr) a_a(u) + \frac{1}{2} r^2 \dot{a}_a(u) + \frac{1}{2} r^2 R_{a0b0}(u)\Omega^b + O(r^3)\Bigr]\, e^\alpha_0.
\end{aligned} \tag{11.8}$$

We now turn to the decomposition of (e^α_0, e^α_a) in the tetrad $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ associated with the simultaneous point $z(t)$. This decomposition will be expressed in the Fermi normal coordinates as an expansion in powers of s. Here, as in Sec. 9.2, we shall express quantities at $z(u)$ in terms of quantities at $z(t)$. We begin with

$$e^\alpha_0 = p^\alpha(t) - \dot{p}^\alpha(t)\Delta + \frac{1}{2}\ddot{p}^\alpha(t)\Delta^2 + O(\Delta^3)$$

and we evaluate the derivatives of $p^\alpha(\tau)$ at $\tau = t$. To accomplish this we rely on our previous results (replacing primed indices with barred indices), on the expansions of Eq. (6.11), and on the decomposition of $g^\alpha_{\ \bar{\alpha}}(x, \bar{x})$ in the tetrads at x and $\bar{x}$. This gives

$$\begin{aligned}
\dot{p}^\alpha(t) &= \Bigl[a^a + \frac{1}{2} s R^a_{\ 0b0}\omega^b + O(s^2)\Bigr]\bar{e}^\alpha_a, \\
\ddot{p}^\alpha(t) &= \Bigl[-\dot{a}_0 + O(s)\Bigr]\bar{e}^\alpha_0 + \Bigl[\dot{a}^a + O(s)\Bigr]\bar{e}^\alpha_a,
\end{aligned}$$

and we finally obtain

$$e^\alpha_0 = \Bigl[1 - \frac{1}{2} s^2 \dot{a}_0(t) + O(s^3)\Bigr]\,\bar{e}^\alpha_0 + \Bigl[-s\Bigl(1 - \frac{1}{2} s a_b \omega^b\Bigr) a^a(t) + \frac{1}{2} s^2 \dot{a}^a(t) - \frac{1}{2} s^2 R^a_{\ 0b0}(t)\omega^b + O(s^3)\Bigr]\,\bar{e}^\alpha_a. \tag{11.9}$$

Similarly, we write

$$e^\alpha_a = p^\alpha_a(t) - \dot{p}^\alpha_a(t)\Delta + \frac{1}{2}\ddot{p}^\alpha_a(t)\Delta^2 + O(\Delta^3),$$

in which we substitute

$$\begin{aligned}
\dot{p}^\alpha_a(t) &= \Bigl[a_a + \frac{1}{2} s R_{a0b0}\omega^b + O(s^2)\Bigr]\bar{e}^\alpha_0 + \Bigl[-\frac{1}{2} s R^b_{\ a0c}\omega^c + O(s^2)\Bigr]\bar{e}^\alpha_b, \\
\ddot{p}^\alpha_a(t) &= \Bigl[\dot{a}_a + O(s)\Bigr]\bar{e}^\alpha_0 + \Bigl[a_a a^b + O(s)\Bigr]\bar{e}^\alpha_b,
\end{aligned}$$

as well as Eq. (11.4) for $\Delta := t - u$. Our final result is

$$\begin{aligned}
e^\alpha_a &= \Bigl[\delta^b_{\ a} + \frac{1}{2} s^2 a^b(t) a_a(t) + \frac{1}{2} s^2 R^b_{\ a0c}(t)\omega^c + O(s^3)\Bigr]\,\bar{e}^\alpha_b \\
&\quad + \Bigl[-s\Bigl(1 - \frac{1}{2} s a_b \omega^b\Bigr) a_a(t) + \frac{1}{2} s^2 \dot{a}_a(t) - \frac{1}{2} s^2 R_{a0b0}(u)\omega^b + O(s^3)\Bigr]\,\bar{e}^\alpha_0.
\end{aligned} \tag{11.10}$$

11.4 Advanced point

It will prove convenient to introduce on the world line, along with the retarded and simultaneous points, an *advanced point* associated with the field point x. The advanced point will be denoted $x'' := z(v)$, with v denoting the value of the proper-time parameter at x''; to tensors at this point we assign indices α'', β'', etc. The advanced point is linked to x by a *past-directed null geodesic* (refer back to Fig. 8), and it can be located by solving $\sigma(x, x'') = 0$ together with the requirement that $\sigma^{\alpha''}(x, x'')$ be a future-directed null vector. The affine-parameter distance between x and x'' along the null geodesic is given by

$$r_{\rm adv} = -\sigma_{\alpha''} u^{\alpha''}, \tag{11.11}$$

and we shall call this the *advanced distance* between x and the world line. Notice that $r_{\rm adv}$ is a positive quantity.

We wish first to find an expression for v in terms of the retarded coordinates of x. For this purpose we define $\Delta' := v - u$ and re-introduce the function $\sigma(\tau) := \sigma(x, z(\tau))$ first considered in Sec. 11.2. We have that $\sigma(v) = \sigma(u) = 0$, and Δ' can ultimately be obtained by expressing $\sigma(v)$ as $\sigma(u + \Delta')$ and expanding in powers of Δ'. Recalling that $\dot\sigma(\tau) = p(\tau)$, we have

$$\sigma(v) = \sigma(u) + p(u)\Delta' + \frac{1}{2}\dot p(u)\Delta'^2 + \frac{1}{6}\ddot p(u)\Delta'^3 + \frac{1}{24}p^{(3)}(u)\Delta'^4 + O(\Delta'^5).$$

Using the expressions for the derivatives of $p(\tau)$ that were first obtained in Sec. 11.1, we write this as

$$r = \frac{1}{2}\Big[1 + r a_a \Omega^a + \frac{1}{3}r^2 S + O(r^3)\Big]\Delta' + \frac{1}{6}r\Big[\dot a_0 + \dot a_a\Omega^a + O(r)\Big]\Delta'^2 - \frac{1}{24}\Big[\dot a_0 + O(r)\Big]\Delta'^3 + O(\Delta'^4).$$

Solving for Δ' as an expansion in powers of r, we obtain

$$v = u + 2r\bigg\{1 - r a_a(u)\Omega^a + r^2\big[a_a(u)\Omega^a\big]^2 - \frac{1}{3}r^2\dot a_0(u) - \frac{2}{3}r^2\dot a_a(u)\Omega^a - \frac{1}{3}r^2 R_{a0b0}(u)\Omega^a\Omega^b + O(r^3)\bigg\}, \tag{11.12}$$

in which all frame components are evaluated at the retarded point $z(u)$.

Our next task is to derive an expression for the advanced distance $r_{\rm adv}$. For this purpose we observe that $r_{\rm adv} = -p(v) = -p(u + \Delta')$, which we can expand in powers of $\Delta' := v - u$. This gives

$$r_{\rm adv} = -p(u) - \dot p(u)\Delta' - \frac{1}{2}\ddot p(u)\Delta'^2 - \frac{1}{6}p^{(3)}(u)\Delta'^3 + O(\Delta'^4),$$

which then becomes

$$r_{\rm adv} = -r + \Big[1 + r a_a\Omega^a + \frac{1}{3}r^2 S + O(r^3)\Big]\Delta' + \frac{1}{2}r\Big[\dot a_0 + \dot a_a\Omega^a + O(r)\Big]\Delta'^2 - \frac{1}{6}\Big[\dot a_0 + O(r)\Big]\Delta'^3 + O(\Delta'^4).$$

After substituting Eq. (11.12) for Δ' and witnessing a number of cancellations, we arrive at the simple expression

$$r_{\rm adv} = r\Big[1 + \frac{2}{3}r^2\dot a_a(u)\Omega^a + O(r^3)\Big]. \tag{11.13}$$

From Eqs. (10.29), (10.30), and (11.12) we deduce that the gradient of the advanced time v is given by

$$\partial_\alpha v = \Big[1 - 2r a_a\Omega^a + O(r^2)\Big]\, e^0_\alpha + \Big[\Omega_a - 2r a_a + O(r^2)\Big]\, e^a_\alpha, \tag{11.14}$$

where the expansion in powers of r was truncated to a sufficient number of terms. Similarly, Eqs. (10.30), (10.31), and (11.13) imply that the gradient of the advanced distance is given by

$$\begin{aligned}\partial_\alpha r_{\rm adv} \;&=\; \Big[\Big(1 + r a_b\Omega^b + \frac{4}{3}r^2\dot a_b\Omega^b + \frac{1}{3}r^2 S\Big)\Omega_a + \frac{2}{3}r^2\dot a_a + \frac{1}{6}r^2 S_a + O(r^3)\Big]\, e^a_\alpha \\ &\quad + \Big[-r a_a\Omega^a - \frac{1}{2}r^2 S + O(r^3)\Big]\, e^0_\alpha,\end{aligned} \tag{11.15}$$

where S_a and S were first introduced in Eqs. (10.24) and (10.25), respectively. We emphasize that in Eqs. (11.14) and (11.15), all frame components are evaluated at the retarded point $z(u)$.

Part III
Green's functions

12 Scalar Green's functions in flat spacetime

12.1 Green's equation for a massive scalar field

To prepare the way for our discussion of Green's functions in curved spacetime, we consider first the slightly nontrivial case of a massive scalar field $\Phi(x)$ in flat spacetime. This field satisfies the wave equation

$$(\Box - k^2)\Phi(x) = -4\pi\mu(x), \tag{12.1}$$

where $\Box = \eta^{\alpha\beta}\partial_\alpha\partial_\beta$ is the wave operator, $\mu(x)$ a prescribed source, and where the parameter k has a dimension of inverse length. We seek a Green's function $G(x,x')$ such that a solution to Eq. (12.1) can be expressed as

$$\Phi(x) = \int G(x,x')\mu(x')\,d^4x', \tag{12.2}$$

where the integration is over all of Minkowski spacetime. The relevant wave equation for the Green's function is

$$(\Box - k^2)G(x,x') = -4\pi\delta_4(x-x'), \tag{12.3}$$

where $\delta_4(x-x') = \delta(t-t')\delta(x-x')\delta(y-y')\delta(z-z')$ is a four-dimensional Dirac distribution in flat spacetime. Two types of Green's functions will be of particular interest: the retarded Green's function, a solution to Eq. (12.3) with the property that it vanishes when x is in the past of x', and the advanced Green's function, which vanishes when x is in the future of x'.

To solve Eq. (12.3) we appeal to Lorentz invariance and the fact that the spacetime is homogeneous to argue that the retarded and advanced Green's functions must be given by expressions of the form

$$G_{\rm ret}(x,x') = \theta(t-t')g(\sigma), \qquad G_{\rm adv}(x,x') = \theta(t'-t)g(\sigma), \tag{12.4}$$

where $\sigma = \frac{1}{2}\eta_{\alpha\beta}(x-x')^\alpha(x-x')^\beta$ is Synge's world function in flat spacetime, and where $g(\sigma)$ is a function to be determined. For the remainder of this section we set $x'=0$ without loss of generality.

12.2 Integration over the source

The Dirac functional on the right-hand side of Eq. (12.3) is a highly singular quantity, and we can avoid dealing with it by integrating the equation over a small four-volume V that contains $x' \equiv 0$. This volume is bounded by a closed hypersurface ∂V. After using Gauss' theorem on the first term of Eq. (12.3), we obtain $\oint_{\partial V} G^{;\alpha} d\Sigma_\alpha - k^2\int_V G\,dV = -4\pi$, where $d\Sigma_\alpha$ is a surface element on ∂V. Assuming that the integral of G over V goes to zero in the limit $V \to 0$, we have

$$\lim_{V\to 0}\oint_{\partial V} G^{;\alpha} d\Sigma_\alpha = -4\pi. \tag{12.5}$$

It should be emphasized that the four-volume V must contain the point x'.

To examine Eq. (12.5) we introduce coordinates (w,χ,θ,ϕ) defined by

$$t = w\cos\chi, \qquad x = w\sin\chi\sin\theta\cos\phi, \qquad y = w\sin\chi\sin\theta\sin\phi, \qquad z = w\sin\chi\cos\theta,$$

and we let ∂V be a surface of constant w. The metric of flat spacetime is given by

$$ds^2 = -\cos 2\chi\,dw^2 + 2w\sin 2\chi\,dwd\chi + w^2\cos 2\chi\,d\chi^2 + w^2\sin^2\chi\,d\Omega^2$$

in the new coordinates, where $d\Omega^2 = d\theta^2 + \sin^2\theta\, d\phi^2$. Notice that w is a timelike coordinate when $\cos 2\chi > 0$, and that χ is then a spacelike coordinate; the roles are reversed when $\cos 2\chi < 0$. Straightforward computations reveal that in these coordinates, $\sigma = -\frac{1}{2}w^2\cos 2\chi$, $\sqrt{-g} = w^3\sin^2\chi\sin\theta$, $g^{ww} = -\cos 2\chi$, $g^{w\chi} = w^{-1}\sin 2\chi$, $g^{\chi\chi} = w^{-2}\cos 2\chi$, and the only nonvanishing component of the surface element is $d\Sigma_w = w^3\sin^2\chi\, d\chi d\Omega$, where $d\Omega = \sin\theta\, d\theta d\phi$. To calculate the gradient of the Green's function we express it as $G = \theta(\pm t)g(\sigma) = \theta(\pm w\cos\chi)g(-\frac{1}{2}w^2\cos 2\chi)$, with the upper (lower) sign belonging to the retarded (advanced) Green's function. Calculation gives $G^{;\alpha}d\Sigma_\alpha = \theta(\pm\cos\chi)w^4\sin^2\chi g'(\sigma)\, d\chi d\Omega$, with a prime indicating differentiation with respect to σ; it should be noted that derivatives of the step function do not appear in this expression.

Integration of $G^{;\alpha}d\Sigma_\alpha$ with respect to $d\Omega$ is immediate, and we find that Eq. (12.5) reduces to

$$\lim_{w\to 0}\int_0^\pi \theta(\pm\cos\chi)w^4\sin^2\chi g'(\sigma)\, d\chi = -1. \tag{12.6}$$

For the retarded Green's function, the step function restricts the domain of integration to $0 < \chi < \pi/2$, in which σ increases from $-\frac{1}{2}w^2$ to $\frac{1}{2}w^2$. Changing the variable of integration from χ to σ transforms Eq. (12.6) into

$$\lim_{\epsilon\to 0}\epsilon\int_{-\epsilon}^{\epsilon} w(\sigma/\epsilon)\, g'(\sigma)\, d\sigma = -1, \qquad w(\xi) := \sqrt{\frac{1+\xi}{1-\xi}}, \tag{12.7}$$

where $\epsilon := \frac{1}{2}w^2$. For the advanced Green's function, the domain of integration is $\pi/2 < \chi < \pi$, in which σ decreases from $\frac{1}{2}w^2$ to $-\frac{1}{2}w^2$. Changing the variable of integration from χ to σ also produces Eq. (12.7).

12.3 Singular part of $g(\sigma)$

We have seen that Eq. (12.7) properly encodes the influence of the singular source term on both the retarded and advanced Green's function. The function $g(\sigma)$ that enters into the expressions of Eq. (12.4) must therefore be such that Eq. (12.7) is satisfied. It follows immediately that $g(\sigma)$ must be a singular function, because for a smooth function the integral of Eq. (12.7) would be of order ϵ and the left-hand side of Eq. (12.7) could never be made equal to -1. The singularity, however, must be integrable, and this leads us to assume that $g'(\sigma)$ must be made out of Dirac δ-functions and derivatives.

We make the ansatz

$$g(\sigma) = V(\sigma)\theta(-\sigma) + A\delta(\sigma) + B\delta'(\sigma) + C\delta''(\sigma) + \cdots, \tag{12.8}$$

where $V(\sigma)$ is a smooth function, and A, B, C, …are constants. The first term represents a function supported within the past and future light cones of $x' \equiv 0$; we exclude a term proportional to $\theta(\sigma)$ for reasons of causality. The other terms are supported on the past and future light cones. It is sufficient to take the coefficients in front of the δ-functions to be constants. To see this we invoke the distributional identities

$$\sigma\delta(\sigma) = 0 \quad\rightarrow\quad \sigma\delta'(\sigma) + \delta(\sigma) = 0 \quad\rightarrow\quad \sigma\delta''(\sigma) + 2\delta'(\sigma) = 0 \quad\rightarrow\quad \cdots \tag{12.9}$$

from which it follows that $\sigma^2\delta'(\sigma) = \sigma^3\delta''(\sigma) = \cdots = 0$. A term like $f(\sigma)\delta(\sigma)$ is then distributionally equal to $f(0)\delta(\sigma)$, while a term like $f(\sigma)\delta'(\sigma)$ is distributionally equal to $f(0)\delta'(\sigma) - f'(0)\delta(\sigma)$, and a term like $f(\sigma)\delta''(\sigma)$ is distributionally equal to $f(0)\delta''(\sigma) - 2f'(0)\delta'(\sigma) + 2f''(0)\delta(\sigma)$; here $f(\sigma)$ is an arbitrary test function. Summing over such terms, we recover an expression of the form of Eq. (12.9), and there is no need to make A, B, C, …functions of σ.

Differentiation of Eq. (12.8) and substitution into Eq. (12.7) yields

$$\epsilon\int_{-\epsilon}^{\epsilon} w(\sigma/\epsilon)\, g'(\sigma)\, d\sigma = \epsilon\left[\int_{-\epsilon}^{\epsilon} V'(\sigma)w(\sigma/\epsilon)\, d\sigma - V(0)w(0) - \frac{A}{\epsilon}\dot{w}(0) + \frac{B}{\epsilon^2}\ddot{w}(0) - \frac{C}{\epsilon^3}w^{(3)}(0) + \cdots\right],$$

where overdots (or a number within brackets) indicate repeated differentiation with respect to $\xi := \sigma/\epsilon$. The limit $\epsilon \to 0$ exists if and only if $B = C = \cdots = 0$. In the limit we must then have $A\dot{w}(0) = 1$, which implies $A = 1$. We conclude that $g(\sigma)$ must have the form of

$$g(\sigma) = \delta(\sigma) + V(\sigma)\theta(-\sigma), \tag{12.10}$$

with $V(\sigma)$ a smooth function that cannot be determined from Eq. (12.7) alone.

12.4 Smooth part of $g(\sigma)$

To determine $V(\sigma)$ we must go back to the differential equation of Eq. (12.3). Because the singular structure of the Green's function is now under control, we can safely set $x \neq x' \equiv 0$ in the forthcoming operations. This means that the equation to solve is in fact $(\Box - k^2)g(\sigma) = 0$, the homogeneous version of Eq. (12.3). We have $\nabla_\alpha g = g'\sigma_\alpha$, $\nabla_\alpha\nabla_\beta g = g''\sigma_\alpha\sigma_\beta + g'\sigma_{\alpha\beta}$, $\Box g = 2\sigma g'' + 4g'$, so that Green's equation reduces to the ordinary differential equation

$$2\sigma g'' + 4g' - k^2 g = 0. \tag{12.11}$$

If we substitute Eq. (12.10) into this we get

$$-(2V + k^2)\delta(\sigma) + (2\sigma V'' + 4V' - k^2 V)\theta(-\sigma) = 0,$$

where we have used the identities of Eq. (12.9). The left-hand side will vanish as a distribution if we set

$$2\sigma V'' + 4V' - k^2 V = 0, \qquad V(0) = -\frac{1}{2}k^2. \tag{12.12}$$

These equations determine $V(\sigma)$ uniquely, even in the absence of a second boundary condition at $\sigma = 0$, because the differential equation is singular at $\sigma = 0$ while V is known to be smooth.

To solve Eq. (12.12) we let $V = F(z)/z$, with $z := k\sqrt{-2\sigma}$. This gives rise to Bessel's equation for the new function F:

$$z^2 F_{zz} + zF_z + (z^2 - 1)F = 0.$$

The solution that is well behaved near $z = 0$ is $F = aJ_1(z)$, where a is a constant to be determined. We have that $J_1(z) \sim \frac{1}{2}z$ for small values of z, and it follows that $V \sim a/2$. From Eq. (12.12) we see that $a = -k^2$. So we have found that the only acceptable solution to Eq. (12.12) is

$$V(\sigma) = -\frac{k}{\sqrt{-2\sigma}}\, J_1\bigl(k\sqrt{-2\sigma}\bigr). \tag{12.13}$$

To summarize, the retarded and advanced solutions to Eq. (12.3) are given by Eq. (12.4) with $g(\sigma)$ given by Eq. (12.10) and $V(\sigma)$ given by Eq. (12.13).

12.5 Advanced distributional methods

The techniques developed previously to find Green's functions for the scalar wave equation are limited to flat spacetime, and they would not be very useful for curved spacetimes. To pursue this generalization we must introduce more powerful distributional methods. We do so in this subsection, and in the next we shall use them to recover our previous results.

Let $\theta_+(x, \Sigma)$ be a generalized step function, defined to be one when x is in the future of the spacelike hypersurface Σ and zero otherwise. Similarly, define $\theta_-(x, \Sigma) := 1 - \theta_+(x, \Sigma)$ to be one when x is in the past of the spacelike hypersurface Σ and zero otherwise. Then define the light-cone step functions

$$\theta_\pm(-\sigma) = \theta_\pm(x, \Sigma)\theta(-\sigma), \qquad x' \in \Sigma, \tag{12.14}$$

so that $\theta_+(-\sigma)$ is one if x is within $I^+(x')$, the chronological future of x', and zero otherwise, and $\theta_-(-\sigma)$ is one if x is within $I^-(x')$, the chronological past of x', and zero otherwise; the choice of hypersurface is immaterial so long as Σ is spacelike and contains the reference point x'. Notice that $\theta_+(-\sigma) + \theta_-(-\sigma) = \theta(-\sigma)$. Define also the light-cone Dirac functionals

$$\delta_\pm(\sigma) = \theta_\pm(x, \Sigma)\delta(\sigma), \qquad x' \in \Sigma, \tag{12.15}$$

so that $\delta_+(\sigma)$, when viewed as a function of x, is supported on the future light cone of x', while $\delta_-(\sigma)$ is supported on its past light cone. Notice that $\delta_+(\sigma) + \delta_-(\sigma) = \delta(\sigma)$. In Eqs. (12.14) and (12.15), σ is the world function for flat spacetime; it is negative when x and x' are timelike related, and positive when they are spacelike related.

The distributions $\theta_\pm(-\sigma)$ and $\delta_\pm(\sigma)$ are not defined at $x = x'$ and they cannot be differentiated there. This pathology can be avoided if we shift σ by a small positive quantity ϵ. We can therefore use the

distributions $\theta_\pm(-\sigma-\epsilon)$ and $\delta_\pm(\sigma+\epsilon)$ in some sensitive computations, and then take the limit $\epsilon\to 0^+$. Notice that the equation $\sigma+\epsilon=0$ describes a two-branch hyperboloid that is located just *within* the light cone of the reference point x'. The hyperboloid does not include x', and $\theta_+(x,\Sigma)$ is one everywhere on its future branch, while $\theta_-(x,\Sigma)$ is one everywhere on its past branch. These factors, therefore, become invisible to differential operators. For example, $\theta_+'(-\sigma-\epsilon)=\theta_+(x,\Sigma)\theta'(-\sigma-\epsilon)=-\theta_+(x,\Sigma)\delta(\sigma+\epsilon)=-\delta_+(\sigma+\epsilon)$. This manipulation shows that after the shift from σ to $\sigma+\epsilon$, the distributions of Eqs. (12.14) and (12.15) can be straightforwardly differentiated with respect to σ.

In the next paragraphs we shall establish the distributional identities

$$\lim_{\epsilon\to 0^+}\epsilon\delta_\pm(\sigma+\epsilon) = 0, \tag{12.16}$$

$$\lim_{\epsilon\to 0^+}\epsilon\delta_\pm'(\sigma+\epsilon) = 0, \tag{12.17}$$

$$\lim_{\epsilon\to 0^+}\epsilon\delta_\pm''(\sigma+\epsilon) = 2\pi\delta_4(x-x') \tag{12.18}$$

in four-dimensional flat spacetime. These will be used in the next subsection to recover the Green's functions for the scalar wave equation, and they will be generalized to curved spacetime in Sec. 13.

The derivation of Eqs. (12.16)–(12.18) relies on a "master" distributional identity, formulated in three-dimensional flat space:

$$\lim_{\epsilon\to 0^+}\frac{\epsilon}{R^5}=\frac{2\pi}{3}\delta_3(\boldsymbol{x}), \qquad R:=\sqrt{r^2+2\epsilon}, \tag{12.19}$$

with $r:=|\boldsymbol{x}|:=\sqrt{x^2+y^2+z^2}$. This follows from yet another identity, $\nabla^2 r^{-1}=-4\pi\delta_3(\boldsymbol{x})$, in which we write the left-hand side as $\lim_{\epsilon\to 0^+}\nabla^2 R^{-1}$; since R^{-1} is nonsingular at $\boldsymbol{x}=0$ it can be straightforwardly differentiated, and the result is $\nabla^2 R^{-1}=-6\epsilon/R^5$, from which Eq. (12.19) follows.

To prove Eq. (12.16) we must show that $\epsilon\delta_\pm(\sigma+\epsilon)$ vanishes as a distribution in the limit $\epsilon\to 0^+$. For this we must prove that a functional of the form

$$A_\pm[f]=\lim_{\epsilon\to 0^+}\int \epsilon\delta_\pm(\sigma+\epsilon)f(x)\,d^4x,$$

where $f(x)=f(t,\boldsymbol{x})$ is a smooth test function, vanishes for all such functions f. Our first task will be to find a more convenient expression for $\delta_\pm(\sigma+\epsilon)$. Once more we set $x'=0$ (without loss of generality) and we note that $2(\sigma+\epsilon)=-t^2+r^2+2\epsilon=-(t-R)(t+R)$, where we have used Eq. (12.19). It follows that

$$\delta_\pm(\sigma+\epsilon)=\frac{\delta(t\mp R)}{R}, \tag{12.20}$$

and from this we find

$$A_\pm[f]=\lim_{\epsilon\to 0^+}\int \epsilon\frac{f(\pm R,\boldsymbol{x})}{R}\,d^3x=\lim_{\epsilon\to 0^+}\int\frac{\epsilon}{R^5}R^4 f(\pm R,\boldsymbol{x})\,d^3x=\frac{2\pi}{3}\int\delta_3(\boldsymbol{x})r^4 f(\pm r,\boldsymbol{x})\,d^3x=0,$$

which establishes Eq. (12.16).

The validity of Eq. (12.17) is established by a similar computation. Here we must show that a functional of the form

$$B_\pm[f]=\lim_{\epsilon\to 0^+}\int\epsilon\delta_\pm'(\sigma+\epsilon)f(x)\,d^4x$$

vanishes for all test functions f. We have

$$\begin{aligned}B_\pm[f] &= \lim_{\epsilon\to 0^+}\epsilon\frac{d}{d\epsilon}\int\delta_\pm(\sigma+\epsilon)f(x)\,d^4x=\lim_{\epsilon\to 0^+}\epsilon\frac{d}{d\epsilon}\int\frac{f(\pm R,\boldsymbol{x})}{R}\,d^3x=\lim_{\epsilon\to 0^+}\epsilon\int\biggl(\pm\frac{\dot f}{R^2}-\frac{f}{R^3}\biggr)\,d^3x\\ &= \lim_{\epsilon\to 0^+}\int\frac{\epsilon}{R^5}\bigl(\pm R^3\dot f-R^2 f\bigr)\,d^3x=\frac{2\pi}{3}\int\delta_3(\boldsymbol{x})\bigl(\pm r^3\dot f-r^2 f\bigr)\,d^3x=0,\end{aligned}$$

and the identity of Eq. (12.17) is proved. In these manipulations we have let an overdot indicate partial differentiation with respect to t, and we have used $\partial R/\partial\epsilon=1/R$.

To establish Eq. (12.18) we consider the functional

$$C_{\pm}[f] = \lim_{\epsilon \to 0^+} \int \epsilon \delta''_{\pm}(\sigma + \epsilon) f(x)\, d^4x$$

and show that it evaluates to $2\pi f(0, \mathbf{0})$. We have

$$\begin{aligned} C_{\pm}[f] &= \lim_{\epsilon \to 0^+} \epsilon \frac{d^2}{d\epsilon^2} \int \delta_{\pm}(\sigma + \epsilon) f(x)\, d^4x = \lim_{\epsilon \to 0^+} \epsilon \frac{d^2}{d\epsilon^2} \int \frac{f(\pm R, \boldsymbol{x})}{R}\, d^3x \\ &= \lim_{\epsilon \to 0^+} \epsilon \int \biggl(\frac{\ddot{f}}{R^3} \mp 3 \frac{\dot{f}}{R^4} + 3 \frac{f}{R^5} \biggr) d^3x = 2\pi \int \delta_3(\boldsymbol{x}) \biggl(\frac{1}{3} r^2 \ddot{f} \pm r \dot{f} + f \biggr) d^3x \\ &= 2\pi f(0, \mathbf{0}), \end{aligned}$$

as required. This proves that Eq. (12.18) holds as a distributional identity in four-dimensional flat spacetime.

12.6 Alternative computation of the Green's functions

The retarded and advanced Green's functions for the scalar wave equation are now defined as the limit of the functions $G^{\epsilon}_{\pm}(x, x')$ when $\epsilon \to 0^+$. For these we make the ansatz

$$G^{\epsilon}_{\pm}(x, x') = \delta_{\pm}(\sigma + \epsilon) + V(\sigma)\theta_{\pm}(-\sigma - \epsilon), \tag{12.21}$$

and we shall prove that $G^{\epsilon}_{\pm}(x, x')$ satisfies Eq. (12.3) in the limit. We recall that the distributions $\theta_{\pm}$ and $\delta_{\pm}$ were defined in the preceding subsection, and we assume that $V(\sigma)$ is a smooth function of $\sigma(x, x') = \frac{1}{2}\eta_{\alpha\beta}(x - x')^{\alpha}(x - x')^{\beta}$; because this function is smooth, it is not necessary to evaluate V at $\sigma + \epsilon$ in Eq. (12.21). We recall also that θ_+ and δ_+ are nonzero when x is in the future of x', while θ_- and δ_- are nonzero when x is in the past of x'. We will therefore prove that the retarded and advanced Green's functions are of the form

$$G_{\rm ret}(x, x') = \lim_{\epsilon \to 0^+} G^{\epsilon}_{+}(x, x') = \theta_+(x, \Sigma)\bigl[\delta(\sigma) + V(\sigma)\theta(-\sigma)\bigr] \tag{12.22}$$

and

$$G_{\rm adv}(x, x') = \lim_{\epsilon \to 0^+} G^{\epsilon}_{-}(x, x') = \theta_-(x, \Sigma)\bigl[\delta(\sigma) + V(\sigma)\theta(-\sigma)\bigr], \tag{12.23}$$

where Σ is a spacelike hypersurface that contains x'. We will also determine the form of the function $V(\sigma)$.

The functions that appear in Eq. (12.21) can be straightforwardly differentiated. The manipulations are similar to what was done in Sec. 12.4, and dropping all labels, we obtain $(\Box - k^2)G = 2\sigma G'' + 4G' - k^2 G$, with a prime indicating differentiation with respect to σ. From Eq. (12.21) we obtain $G' = \delta' - V\delta + V'\theta$ and $G'' = \delta'' - V\delta' - 2V'\delta + V''\theta$. The identities of Eq. (12.9) can be expressed as $(\sigma + \epsilon)\delta'(\sigma + \epsilon) = -\delta(\sigma + \epsilon)$ and $(\sigma + \epsilon)\delta''(\sigma + \epsilon) = -2\delta'(\sigma + \epsilon)$, and combining this with our previous results gives

$$\begin{aligned} (\Box - k^2)G^{\epsilon}_{\pm}(x, x') &= (-2V - k^2)\delta_{\pm}(\sigma + \epsilon) + (2\sigma V'' + 4V' - k^2 V)\theta_{\pm}(-\sigma - \epsilon) \\ &\quad - 2\epsilon\delta''_{\pm}(\sigma + \epsilon) + 2V\epsilon\delta'_{\pm}(\sigma + \epsilon) + 4V'\epsilon\delta_{\pm}(\sigma + \epsilon). \end{aligned}$$

According to Eq. (12.16)–(12.18), the last two terms on the right-hand side disappear in the limit $\epsilon \to 0^+$, and the third term becomes $-4\pi\delta_4(x - x')$. Provided that the first two terms vanish also, we recover $(\Box - k^2)G(x, x') = -4\pi\delta_4(x - x')$ in the limit, as required. Thus, the limit of $G^{\epsilon}_{\pm}(x, x')$ when $\epsilon \to 0^+$ will indeed satisfy Green's equation provided that $V(\sigma)$ is a solution to

$$2\sigma V'' + 4V' - k^2 V = 0, \qquad V(0) = -\frac{1}{2}k^2; \tag{12.24}$$

these are the same statements as in Eq. (12.12). The solution to these equations was produced in Eq. (12.13):

$$V(\sigma) = -\frac{k}{\sqrt{-2\sigma}}\, J_1\bigl(k\sqrt{-2\sigma}\bigr), \tag{12.25}$$

and this completely determines the Green's functions of Eqs. (12.22) and (12.23).

13 Distributions in curved spacetime

The distributions introduced in Sec. 12.5 can also be defined in a four-dimensional spacetime with metric $g_{\alpha\beta}$. Here we produce the relevant generalizations of the results derived in that section.

13.1 Invariant Dirac distribution

We first introduce $\delta_4(x,x')$, an *invariant* Dirac functional in a four-dimensional curved spacetime. This is defined by the relations

$$\int_V f(x)\delta_4(x,x')\sqrt{-g}\,d^4x = f(x'), \qquad \int_{V'} f(x')\delta_4(x,x')\sqrt{-g'}\,d^4x' = f(x), \tag{13.1}$$

where $f(x)$ is a smooth test function, V any four-dimensional region that contains x', and V' any four-dimensional region that contains x. These relations imply that $\delta_4(x,x')$ is symmetric in its arguments, and it is easy to see that

$$\delta_4(x,x') = \frac{\delta_4(x-x')}{\sqrt{-g}} = \frac{\delta_4(x-x')}{\sqrt{-g'}} = (gg')^{-1/4}\delta_4(x-x'), \tag{13.2}$$

where $\delta_4(x-x') = \delta(x^0-x'^0)\delta(x^1-x'^1)\delta(x^2-x'^2)\delta(x^3-x'^3)$ is the ordinary (coordinate) four-dimensional Dirac functional. The relations of Eq. (13.2) are all equivalent because $f(x)\delta_4(x,x') = f(x')\delta_4(x,x')$ is a distributional identity; the last form is manifestly symmetric in x and x'.

The invariant Dirac distribution satisfies the identities

$$\begin{aligned} \Omega_{\dots}(x,x')\delta_4(x,x') &= \big[\Omega_{\dots}\big]\delta_4(x,x'), \\ \big(g^{\alpha}{}_{\alpha'}(x,x')\delta_4(x,x')\big)_{;\alpha} = -\partial_{\alpha'}\delta_4(x,x'), &\qquad \big(g^{\alpha'}{}_{\alpha}(x',x)\delta_4(x,x')\big)_{;\alpha'} = -\partial_{\alpha}\delta_4(x,x'), \end{aligned} \tag{13.3}$$

where $\Omega_{\dots}(x,x')$ is any bitensor and $g^{\alpha}{}_{\alpha'}(x,x')$, $g^{\alpha'}{}_{\alpha}(x,x')$ are parallel propagators. The first identity follows immediately from the definition of the δ-function. The second and third identities are established by showing that integration against a test function $f(x)$ gives the same result from both sides. For example, the first of the Eqs. (13.1) implies

$$\int_V f(x)\partial_{\alpha'}\delta_4(x,x')\sqrt{-g}\,d^4x = \partial_{\alpha'}f(x'),$$

and on the other hand,

$$-\int_V f(x)\big(g^{\alpha}{}_{\alpha'}\delta_4(x,x')\big)_{;\alpha}\sqrt{-g}\,d^4x = -\oint_{\partial V} f(x)g^{\alpha}{}_{\alpha'}\delta_4(x,x')d\Sigma_\alpha + \big[f_{,\alpha}g^{\alpha}{}_{\alpha'}\big] = \partial_{\alpha'}f(x'),$$

which establishes the second identity of Eq. (13.3). Notice that in these manipulations, the integrations involve *scalar* functions of the coordinates x; the fact that these functions are also vectors with respect to x' does not invalidate the procedure. The third identity of Eq. (13.3) is proved in a similar way.

13.2 Light-cone distributions

For the remainder of Sec. 13 we assume that $x \in \mathcal{N}(x')$, so that a unique geodesic β links these two points. We then let $\sigma(x,x')$ be the curved spacetime world function, and we define light-cone step functions by

$$\theta_\pm(-\sigma) = \theta_\pm(x,\Sigma)\theta(-\sigma), \qquad x' \in \Sigma, \tag{13.4}$$

where $\theta_+(x,\Sigma)$ is one when x is in the future of the spacelike hypersurface Σ and zero otherwise, and $\theta_-(x,\Sigma) = 1-\theta_+(x,\Sigma)$. These are immediate generalizations to curved spacetime of the objects defined in flat spacetime by Eq. (12.14). We have that $\theta_+(-\sigma)$ is one when x is within $I^+(x')$, the chronological future of x', and zero otherwise, and $\theta_-(-\sigma)$ is one when x is within $I^-(x')$, the chronological past of x', and zero otherwise. We also have $\theta_+(-\sigma)+\theta_-(-\sigma) = \theta(-\sigma)$.

We define the curved-spacetime version of the light-cone Dirac functionals by

$$\delta_\pm(\sigma) = \theta_\pm(x,\Sigma)\delta(\sigma), \qquad x' \in \Sigma, \tag{13.5}$$

an immediate generalization of Eq. (12.15). We have that $\delta_+(\sigma)$, when viewed as a function of x, is supported on the future light cone of x', while $\delta_-(\sigma)$ is supported on its past light cone. We also have $\delta_+(\sigma)+\delta_-(\sigma) = \delta(\sigma)$, and we recall that σ is negative when x and x' are timelike related, and positive when they are spacelike related.

For the same reasons as those mentioned in Sec. 12.5, it is sometimes convenient to shift the argument of the step and δ-functions from σ to $\sigma+\epsilon$, where ϵ is a small positive quantity. With this shift, the light-cone distributions can be straightforwardly differentiated with respect to σ. For example, $\delta_\pm(\sigma+\epsilon) = -\theta'_\pm(-\sigma-\epsilon)$, with a prime indicating differentiation with respect to σ.

We now prove that the identities of Eq. (12.16)–(12.18) generalize to

$$\lim_{\epsilon\to 0^+} \epsilon\delta_\pm(\sigma+\epsilon) \quad = \quad 0, \tag{13.6}$$

$$\lim_{\epsilon\to 0^+} \epsilon\delta'_\pm(\sigma+\epsilon) \quad = \quad 0, \tag{13.7}$$

$$\lim_{\epsilon\to 0^+} \epsilon\delta''_\pm(\sigma+\epsilon) \quad = \quad 2\pi\delta_4(x,x') \tag{13.8}$$

in a four-dimensional curved spacetime; the only differences lie with the definition of the world function and the fact that it is the invariant Dirac functional that appears in Eq. (13.8). To establish these identities in curved spacetime we use the fact that they hold in flat spacetime — as was shown in Sec. 12.5 — and that they are scalar relations that must be valid in any coordinate system if they are found to hold in one. Let us then examine Eqs. (13.6)–(13.7) in the Riemann normal coordinates of Sec. 8; these are denoted $\hat{x}^\alpha$ and are based at x'. We have that $\sigma(x,x') = \frac{1}{2}\eta_{\alpha\beta}\hat{x}^\alpha\hat{x}^\beta$ and $\delta_4(x,x') = \Delta(x,x')\delta_4(x-x') = \delta_4(x-x')$, where $\Delta(x,x')$ is the van Vleck determinant, whose coincidence limit is unity. In Riemann normal coordinates, therefore, Eqs. (13.6)–(13.8) take exactly the same form as Eqs. (12.16)–(12.18). Because the identities are true in flat spacetime, they must be true also in curved spacetime (in Riemann normal coordinates based at x'); and because these are scalar relations, they must be valid in any coordinate system.

14 Scalar Green's functions in curved spacetime

14.1 Green's equation for a massless scalar field in curved spacetime

We consider a massless scalar field $\Phi(x)$ in a curved spacetime with metric $g_{\alpha\beta}$. The field satisfies the wave equation

$$(\Box - \xi R)\Phi(x) = -4\pi\mu(x), \tag{14.1}$$

where $\Box = g^{\alpha\beta}\nabla_\alpha\nabla_\beta$ is the wave operator, R the Ricci scalar, ξ an arbitrary coupling constant, and $\mu(x)$ is a prescribed source. We seek a Green's function $G(x,x')$ such that a solution to Eq. (14.1) can be expressed as

$$\Phi(x) = \int G(x,x')\mu(x')\sqrt{-g'}\, d^4x', \tag{14.2}$$

where the integration is over the entire spacetime. The wave equation for the Green's function is

$$(\Box - \xi R)G(x,x') = -4\pi\delta_4(x,x'), \tag{14.3}$$

where $\delta_4(x,x')$ is the invariant Dirac functional introduced in Sec. 13.1. It is easy to verify that the field defined by Eq. (14.2) is truly a solution to Eq. (14.1).

We let $G_+(x,x')$ be the retarded solution to Eq. (14.3), and $G_-(x,x')$ is the advanced solution; when viewed as functions of x, $G_+(x,x')$ is nonzero in the causal future of x', while $G_-(x,x')$ is nonzero in its causal past. We assume that the retarded and advanced Green's functions exist as distributions and can be defined globally in the entire spacetime.

14.2 Hadamard construction of the Green's functions

Assuming throughout this subsection that x is restricted to the normal convex neighbourhood of x', we make the ansatz

$$G_\pm(x,x') = U(x,x')\delta_\pm(\sigma) + V(x,x')\theta_\pm(-\sigma), \tag{14.4}$$

where $U(x,x')$ and $V(x,x')$ are smooth biscalars; the fact that the spacetime is no longer homogeneous means that these functions cannot depend on σ alone.

Before we substitute the Green's functions of Eq. (14.4) into the differential equation of Eq. (14.3) we proceed as in Sec. 12.6 and shift σ by the small positive quantity ϵ. We shall therefore consider the distributions

$$G^\epsilon_\pm(x,x') = U(x,x')\delta_\pm(\sigma+\epsilon) + V(x,x')\theta_\pm(-\theta-\epsilon),$$

and later recover the Green's functions by taking the limit $\epsilon \to 0^+$. Differentiation of these objects is straightforward, and in the following manipulations we will repeatedly use the relation $\sigma^\alpha\sigma_\alpha = 2\sigma$ satisfied by the world function. We will also use the distributional identities $\sigma\delta_\pm(\sigma+\epsilon) = -\epsilon\delta_\pm(\sigma+\epsilon)$, $\sigma\delta'_\pm(\sigma+\epsilon) = -\delta_\pm(\sigma+\epsilon) - \epsilon\delta'_\pm(\sigma+\epsilon)$, and $\sigma\delta''_\pm(\sigma+\epsilon) = -2\delta'(\sigma+\epsilon) - \epsilon\delta''(\sigma+\epsilon)$. After a routine calculation we obtain

$$\begin{aligned}(\Box - \xi R)G^\epsilon_\pm &= -2\epsilon\delta''_\pm(\sigma+\epsilon)U + 2\epsilon\delta'_\pm(\sigma+\epsilon)V + \delta'_\pm(\sigma+\epsilon)\Bigl\{2U_{,\alpha}\sigma^\alpha + (\sigma^\alpha{}_\alpha - 4)U\Bigr\} \\ &\quad + \delta_\pm(\sigma+\epsilon)\Bigl\{-2V_{,\alpha}\sigma^\alpha + (2-\sigma^\alpha{}_\alpha)V + (\Box - \xi R)U\Bigr\} + \theta_\pm(-\sigma-\epsilon)\Bigl\{(\Box-\xi R)V\Bigr\},\end{aligned}$$

which becomes

$$\begin{aligned}(\Box - \xi R)G_\pm &= -4\pi\delta_4(x,x')U + \delta'_\pm(\sigma)\Bigl\{2U_{,\alpha}\sigma^\alpha + (\sigma^\alpha{}_\alpha - 4)U\Bigr\} \\ &\quad + \delta_\pm(\sigma)\Bigl\{-2V_{,\alpha}\sigma^\alpha + (2-\sigma^\alpha{}_\alpha)V + (\Box - \xi R)U\Bigr\} + \theta_\pm(-\sigma)\Bigl\{(\Box-\xi R)V\Bigr\}\end{aligned} \tag{14.5}$$

in the limit $\epsilon \to 0^+$, after using the identities of Eqs. (13.6)–(13.8).

According to Eq. (14.3), the right-hand side of Eq. (14.5) should be equal to $-4\pi\delta_4(x,x')$. This immediately gives us the coincidence condition

$$\bigl[U\bigr] = 1 \tag{14.6}$$

for the biscalar $U(x,x')$. To eliminate the $\delta'_\pm$ term we make its coefficient vanish:

$$2U_{,\alpha}\sigma^\alpha + (\sigma^\alpha{}_\alpha - 4)U = 0. \tag{14.7}$$

As we shall now prove, these two equations determine $U(x,x')$ uniquely.

Recall from Sec. 3.3 that σ^α is a vector at x that is tangent to the unique geodesic β that connects x to x'. This geodesic is affinely parameterized by λ and a displacement along β is described by $dx^\alpha = (\sigma^\alpha/\lambda)d\lambda$. The first term of Eq. (14.7) therefore represents the logarithmic rate of change of $U(x,x')$ along β, and this can be expressed as $2\lambda dU/d\lambda$. For the second term we recall from Sec. 7.1 the differential equation $\Delta^{-1}(\Delta\sigma^\alpha)_{;\alpha} = 4$ satisfied by $\Delta(x,x')$, the van Vleck determinant. This gives us $\sigma^\alpha{}_\alpha - 4 = -\Delta^{-1}\Delta_{,\alpha}\sigma^\alpha = -\Delta^{-1}\lambda d\Delta/d\lambda$, and Eq. (14.7) becomes

$$\lambda\frac{d}{d\lambda}\bigl(2\ln U - \ln\Delta\bigr) = 0.$$

It follows that U^2/Δ is constant on β, and this must therefore be equal to its value at the starting point x': $U^2/\Delta = [U^2/\Delta] = 1$, by virtue of Eq. (14.6) and the property $[\Delta] = 1$ of the van Vleck determinant. Because this statement must be true for all geodesics β that emanate from x', we have found that the unique solution to Eqs. (14.6) and (14.7) is

$$U(x,x') = \Delta^{1/2}(x,x'). \tag{14.8}$$

We must still consider the remaining terms in Eq. (14.5). The $\delta_\pm$ term can be eliminated by demanding that its coefficient vanish when $\sigma = 0$. This, however, does not constrain its value away from the light cone, and we thus obtain information about $V|_{\sigma=0}$ only. Denoting this by $\check{V}(x,x')$ — the restriction of $V(x,x')$ on the light cone $\sigma(x,x') = 0$ — we have

$$\check{V}_{,\alpha}\sigma^\alpha + \frac{1}{2}\bigl(\sigma^\alpha{}_\alpha - 2\bigr)\check{V} = \frac{1}{2}\bigl(\Box - \xi R\bigr)U\Big|_{\sigma=0}, \tag{14.9}$$

where we indicate that the right-hand side also must be restricted to the light cone. The first term of Eq. (14.9) can be expressed as $\lambda d\check{V}/d\lambda$ and this equation can be integrated along any null geodesic that generates the null cone $\sigma(x,x')=0$. For these integrations to be well posed, however, we must provide initial values at $x=x'$. As we shall now see, these can be inferred from Eq. (14.9) and the fact that $V(x,x')$ must be smooth at coincidence.

Equations (7.4) and (14.8) imply that near coincidence, $U(x,x')$ admits the expansion

$$U = 1 + \frac{1}{12}R_{\alpha'\beta'}\sigma^{\alpha'}\sigma^{\beta'} + O(\lambda^3), \tag{14.10}$$

where $R_{\alpha'\beta'}$ is the Ricci tensor at x' and λ is the affine-parameter distance to x (which can be either on or off the light cone). Differentiation of this relation gives

$$U_{,\alpha} = -\frac{1}{6}g^{\alpha'}_{\ \alpha}R_{\alpha'\beta'}\sigma^{\beta'} + O(\lambda^2), \qquad U_{,\alpha'} = \frac{1}{6}R_{\alpha'\beta'}\sigma^{\beta'} + O(\lambda^2), \tag{14.11}$$

and eventually,

$$\bigl[\Box U\bigr] = \frac{1}{6}R(x'). \tag{14.12}$$

Using also $[\sigma^{\alpha}_{\ \alpha}]=4$, we find that the coincidence limit of Eq. (14.9) gives

$$\bigl[V\bigr] = \frac{1}{12}\Bigl(1-6\xi\Bigr)R(x'), \tag{14.13}$$

and this provides the initial values required for the integration of Eq. (14.9) on the null cone.

Equations (14.9) and (14.13) give us a means to construct $\check{V}(x,x')$, the restriction of $V(x,x')$ on the null cone $\sigma(x,x')=0$. These values can then be used as characteristic data for the wave equation

$$(\Box - \xi R)V(x,x') = 0, \tag{14.14}$$

which is obtained by elimination of the $\theta_\pm$ term in Eq. (14.5). While this certainly does not constitute a practical method to compute the biscalar $V(x,x')$, these considerations show that $V(x,x')$ exists and is unique.

To summarize: We have shown that with $U(x,x')$ given by Eq. (14.8) and $V(x,x')$ determined uniquely by the wave equation of Eq. (14.14) and the characteristic data constructed with Eqs. (14.9) and (14.13), the retarded and advanced Green's functions of Eq. (14.4) do indeed satisfy Eq. (14.3). It should be emphasized that the construction provided in this subsection is restricted to $\mathcal{N}(x')$, the normal convex neighbourhood of the reference point x'.

14.3 Reciprocity

We shall now establish the following reciprocity relation between the (globally defined) retarded and advanced Green's functions:

$$G_-(x',x) = G_+(x,x'). \tag{14.15}$$

Before we get to the proof we observe that by virtue of Eq. (14.15), the biscalar $V(x,x')$ must be symmetric in its arguments:

$$V(x',x) = V(x,x'). \tag{14.16}$$

To go from Eq. (14.15) to Eq. (14.16) we simply note that when $x\in\mathcal{N}(x')$ and belongs to $I^+(x')$, then $G_+(x,x')=V(x,x')$ and $G_-(x',x)=V(x',x)$.

To prove the reciprocity relation we invoke the identities

$$G_+(x,x')(\Box-\xi R)G_-(x,x'') = -4\pi G_+(x,x')\delta_4(x,x'')$$

and

$$G_-(x,x'')(\Box-\xi R)G_+(x,x') = -4\pi G_-(x,x'')\delta_4(x,x')$$

and take their difference. On the left-hand side we have

$$G_+(x,x')\Box G_-(x,x'') - G_-(x,x'')\Box G_+(x,x') = \nabla_\alpha\Bigl(G_+(x,x')\nabla^\alpha G_-(x,x'') - G_-(x,x'')\nabla^\alpha G_+(x,x')\Bigr),$$

while the right-hand side gives

$$-4\pi\Bigl(G_+(x,x')\delta_4(x,x'') - G_-(x,x'')\delta_4(x,x')\Bigr).$$

Integrating both sides over a large four-dimensional region V that contains both x' and x'', we obtain

$$\oint_{\partial V}\Bigl(G_+(x,x')\nabla^\alpha G_-(x,x'') - G_-(x,x'')\nabla^\alpha G_+(x,x')\Bigr)\,d\Sigma_\alpha = -4\pi\Bigl(G_+(x'',x') - G_-(x',x'')\Bigr),$$

where ∂V is the boundary of V. Assuming that the Green's functions fall off sufficiently rapidly at infinity (in the limit $\partial V \to \infty$; this statement imposes some restriction on the spacetime's asymptotic structure), we have that the left-hand side of the equation evaluates to zero in the limit. This gives us the statement $G_+(x'',x') = G_-(x',x'')$, which is just Eq. (14.15) with x'' replacing x.

14.4 Kirchhoff representation

Suppose that the values for a scalar field $\Phi(x')$ and its normal derivative $n^{\alpha'}\nabla_{\alpha'}\Phi(x')$ are known on a spacelike hypersurface Σ. Suppose also that the scalar field satisfies the homogeneous wave equation

$$(\Box - \xi R)\Phi(x) = 0. \tag{14.17}$$

Then the value of the field at a point x in the future of Σ is given by Kirchhoff's formula,

$$\Phi(x) = -\frac{1}{4\pi}\int_\Sigma\Bigl(G_+(x,x')\nabla^{\alpha'}\Phi(x') - \Phi(x')\nabla^{\alpha'}G_+(x,x')\Bigr)\,d\Sigma_{\alpha'}, \tag{14.18}$$

where $d\Sigma_{\alpha'}$ is the surface element on Σ. If $n_{\alpha'}$ is the future-directed unit normal, then $d\Sigma_{\alpha'} = -n_{\alpha'}dV$, with dV denoting the invariant volume element on Σ; notice that $d\Sigma_{\alpha'}$ is past directed.

To establish this result we start with the equations

$$G_-(x',x)(\Box' - \xi R')\Phi(x') = 0, \qquad \Phi(x')(\Box' - \xi R')G_-(x',x) = -4\pi\delta_4(x',x)\Phi(x'),$$

in which x and x' refer to arbitrary points in spacetime. Taking their difference gives

$$\nabla_{\alpha'}\Bigl(G_-(x',x)\nabla^{\alpha'}\Phi(x') - \Phi(x')\nabla^{\alpha'}G_-(x',x)\Bigr) = 4\pi\delta_4(x',x)\Phi(x'),$$

and this we integrate over a four-dimensional region V that is bounded in the past by the hypersurface Σ. We suppose that V contains x and we obtain

$$\oint_{\partial V}\Bigl(G_-(x',x)\nabla^{\alpha'}\Phi(x') - \Phi(x')\nabla^{\alpha'}G_-(x',x)\Bigr)\,d\Sigma_{\alpha'} = 4\pi\Phi(x),$$

where $d\Sigma_{\alpha'}$ is the outward-directed surface element on the boundary ∂V. Assuming that the Green's function falls off sufficiently rapidly into the future, we have that the only contribution to the hypersurface integral is the one that comes from Σ. Since the surface element on Σ points in the direction opposite to the outward-directed surface element on ∂V, we must change the sign of the left-hand side to be consistent with the convention adopted previously. With this change we have

$$\Phi(x) = -\frac{1}{4\pi}\oint_{\partial V}\Bigl(G_-(x',x)\nabla^{\alpha'}\Phi(x') - \Phi(x')\nabla^{\alpha'}G_-(x',x)\Bigr)\,d\Sigma_{\alpha'},$$

which is the same statement as Eq. (14.18) if we take into account the reciprocity relation of Eq. (14.15).

14.5 Singular and regular Green's functions

In part IV of this review we will compute the retarded field of a moving scalar charge, and we will analyze its singularity structure near the world line; this will be part of our effort to understand the effect of the field on the particle's motion. The retarded solution to the scalar wave equation is the physically relevant solution because it properly incorporates outgoing-wave boundary conditions at infinity — the advanced solution would come instead with incoming-wave boundary conditions. The retarded field is singular on the world line because a point particle produces a Coulomb field that diverges at the particle's position. In view of this singular behaviour, it is a subtle matter to describe the field's action on the particle, and to formulate meaningful equations of motion.

When facing this problem in flat spacetime (recall the discussion of Sec. 1.3) it is convenient to decompose the retarded Green's function $G_+(x,x')$ into a *singular* Green's function $G_{\rm S}(x,x') := \frac{1}{2}[G_+(x,x')+G_-(x,x')]$ and a *regular* two-point function $G_{\rm R}(x,x') := \frac{1}{2}[G_+(x,x') - G_-(x,x')]$. The singular Green's function takes its name from the fact that it produces a field with the same singularity structure as the retarded solution: the diverging field near the particle is insensitive to the boundary conditions imposed at infinity. We note also that $G_{\rm S}(x,x')$ satisfies the same wave equation as the retarded Green's function (with a Dirac functional as a source), and that by virtue of the reciprocity relations, it is symmetric in its arguments. The regular two-point function, on the other hand, takes its name from the fact that it satisfies the *homogeneous* wave equation, without the Dirac functional on the right-hand side; it produces a field that is regular on the world line of the moving scalar charge. (We reserve the term "Green's function" to a two-point function that satisfies the wave equation with a Dirac distribution on the right-hand side; when the source term is absent, the object is called a "two-point function".)

Because the singular Green's function is symmetric in its argument, it does not distinguish between past and future, and it produces a field that contains equal amounts of outgoing and incoming radiation — the singular solution describes a standing wave at infinity. Removing $G_{\rm S}(x,x')$ from the retarded Green's function will have the effect of removing the singular behaviour of the field *without affecting the motion of the particle*. The motion is not affected because it is intimately tied to the boundary conditions: If the waves are outgoing, the particle loses energy to the radiation and its motion is affected; if the waves are incoming, the particle gains energy from the radiation and its motion is affected differently. With equal amounts of outgoing and incoming radiation, the particle neither loses nor gains energy and its interaction with the scalar field cannot affect its motion. Thus, subtracting $G_{\rm S}(x,x')$ from the retarded Green's function eliminates the singular part of the field without affecting the motion of the scalar charge. The subtraction leaves behind the regular two-point function, which produces a field that is regular on the world line; it is this field that will govern the motion of the particle. The action of this field is well defined, and it properly encodes the outgoing-wave boundary conditions: the particle will lose energy to the radiation.

In this subsection we attempt a decomposition of the curved-spacetime retarded Green's function into singular and regular pieces. The flat-spacetime relations will have to be amended, however, because of the fact that in a curved spacetime, the advanced Green's function is generally nonzero when x' is in the chronological future of x. This implies that the value of the advanced field at x depends on events x' that will unfold *in the future*; this dependence would be inherited by the regular field (which acts on the particle and determines its motion) if the naive definition $G_{\rm R}(x,x') := \frac{1}{2}[G_+(x,x') - G_-(x,x')]$ were to be adopted.

We shall not adopt this definition. Instead, we shall follow Detweiler and Whiting [13] and introduce a singular Green's function with the properties

S1: $G_{\rm S}(x,x')$ satisfies the inhomogeneous scalar wave equation,

$$(\Box - \xi R)G_{\rm S}(x,x') = -4\pi\delta_4(x,x'); \tag{14.19}$$

S2: $G_{\rm S}(x,x')$ is symmetric in its arguments,

$$G_{\rm S}(x',x) = G_{\rm S}(x,x'); \tag{14.20}$$

S3: $G_{\rm S}(x,x')$ vanishes if x is in the chronological past or future of x',

$$G_{\rm S}(x,x') = 0 \qquad \text{when } x \in I^\pm(x'). \tag{14.21}$$

Properties S1 and S2 ensure that the singular Green's function will properly reproduce the singular behaviour of the retarded solution without distinguishing between past and future; and as we shall see, property S3 ensures that the support of the regular two-point function will not include the chronological future of x.

The regular two-point function is then defined by

$$G_{\rm R}(x,x') = G_+(x,x') - G_{\rm S}(x,x'), \tag{14.22}$$

where $G_+(x,x')$ is the retarded Green's function. This comes with the properties

R1: $G_{\rm R}(x,x')$ satisfies the homogeneous wave equation,

$$(\Box - \xi R)G_{\rm R}(x,x') = 0; \tag{14.23}$$

R2: $G_{\rm R}(x,x')$ agrees with the retarded Green's function if x is in the chronological future of x',

$$G_{\rm R}(x,x') = G_+(x,x') \qquad \text{when } x \in I^+(x'); \tag{14.24}$$

R3: $G_{\rm R}(x,x')$ vanishes if x is in the chronological past of x',

$$G_{\rm R}(x,x') = 0 \qquad \text{when } x \in I^-(x'). \tag{14.25}$$

Property R1 follows directly from Eq. (14.22) and property S1 of the singular Green's function. Properties R2 and R3 follow from S3 and the fact that the retarded Green's function vanishes if x is in past of x'. The properties of the regular two-point function ensure that the corresponding regular field will be nonsingular at the world line, and will depend only on the past history of the scalar charge.

We must still show that such singular and regular Green's functions can be constructed. This relies on the existence of a two-point function $H(x,x')$ that would possess the properties

H1: $H(x,x')$ satisfies the homogeneous wave equation,

$$(\Box - \xi R)H(x,x') = 0; \tag{14.26}$$

H2: $H(x,x')$ is symmetric in its arguments,

$$H(x',x) = H(x,x'); \tag{14.27}$$

H3: $H(x,x')$ agrees with the retarded Green's function if x is in the chronological future of x',

$$H(x,x') = G_+(x,x') \qquad \text{when } x \in I^+(x'); \tag{14.28}$$

H4: $H(x,x')$ agrees with the advanced Green's function if x is in the chronological past of x',

$$H(x,x') = G_-(x,x') \qquad \text{when } x \in I^-(x'). \tag{14.29}$$

With a biscalar $H(x,x')$ satisfying these relations, a singular Green's function defined by

$$G_{\rm S}(x,x') = \frac{1}{2}\Big[G_+(x,x') + G_-(x,x') - H(x,x')\Big] \tag{14.30}$$

will satisfy all the properties listed previously: S1 comes as a consequence of H1 and the fact that both the advanced and the retarded Green's functions are solutions to the inhomogeneous wave equation, S2 follows directly from H2 and the definition of Eq. (14.30), and S3 comes as a consequence of H3, H4 and the properties of the retarded and advanced Green's functions.

The question is now: does such a function $H(x,x')$ exist? We will present a plausibility argument for an affirmative answer. Later in this section we will see that $H(x,x')$ is guaranteed to exist in the local convex neighbourhood of x', where it is equal to $V(x,x')$. And in Sec. 14.6 we will see that there exist particular spacetimes for which $H(x,x')$ can be defined globally.

To satisfy all of H1–H4 might seem a tall order, but it should be possible. We first note that property H4 is not independent from the rest: it follows from H2, H3, and the reciprocity relation (14.15) satisfied by the retarded and advanced Green's functions. Let $x \in I^-(x')$, so that $x' \in I^+(x)$. Then $H(x,x') = H(x',x)$ by H2, and by H3 this is equal to $G_+(x',x)$. But by the reciprocity relation this is also equal to $G_-(x,x')$, and we have obtained H4. Alternatively, and this shall be our point of view in the next paragraph, we can think of H3 as following from H2 and H4.

Because $H(x,x')$ satisfies the homogeneous wave equation (property H1), it can be given the Kirkhoff representation of Eq. (14.18): if Σ is a spacelike hypersurface in the past of both x and x', then

$$H(x,x') = -\frac{1}{4\pi}\int_\Sigma \Big(G_+(x,x'')\nabla^{\alpha''} H(x'',x') - H(x'',x')\nabla^{\alpha''} G_+(x,x'')\Big)\, d\Sigma_{\alpha''},$$

where $d\Sigma_{\alpha''}$ is a surface element on Σ. The hypersurface can be partitioned into two segments, $\Sigma^-(x')$ and $\Sigma - \Sigma^-(x')$, with $\Sigma^-(x')$ denoting the intersection of Σ with $I^-(x')$. To enforce H4 it suffices to choose for $H(x,x')$ initial data on $\Sigma^-(x')$ that agree with the initial data for the advanced Green's function; because both functions satisfy the homogeneous wave equation in $I^-(x')$, the agreement will be preserved in the entire domain of dependence of $\Sigma^-(x')$. The data on $\Sigma - \Sigma^-(x')$ is still free, and it should be possible to choose it so as to make $H(x,x')$ symmetric. Assuming that this can be done, we see that H2 is enforced and we conclude that the properties H1, H2, H3, and H4 can all be satisfied.

When x is restricted to the normal convex neighbourhood of x', properties H1–H4 imply that

$$H(x,x') = V(x,x'); \tag{14.31}$$

it should be stressed here that while $H(x,x')$ is assumed to be defined globally in the entire spacetime, the existence of $V(x,x')$ is limited to $\mathcal{N}(x')$. With Eqs. (14.4) and (14.30) we find that the singular Green's function is given explicitly by

$$G_{\rm S}(x,x') = \frac{1}{2}U(x,x')\delta(\sigma) - \frac{1}{2}V(x,x')\theta(\sigma) \tag{14.32}$$

in the normal convex neighbourhood. Equation (14.32) shows very clearly that the singular Green's function does not distinguish between past and future (property S2), and that its support excludes $I^\pm(x')$, in which $\theta(\sigma) = 0$ (property S3). From Eq. (14.22) we get an analogous expression for the regular two-point function:

$$G_{\rm R}(x,x') = \frac{1}{2}U(x,x')\Big[\delta_+(\sigma) - \delta_-(\sigma)\Big] + V(x,x')\Big[\theta_+(-\sigma) + \frac{1}{2}\theta(\sigma)\Big]. \tag{14.33}$$

This reveals directly that the regular two-point function coincides with $G_+(x,x')$ in $I^+(x')$, in which $\theta(\sigma) = 0$ and $\theta_+(-\sigma) = 1$ (property R2), and that its support does not include $I^-(x')$, in which $\theta(\sigma) = \theta_+(-\sigma) = 0$ (property R3).

14.6 Example: Cosmological Green's functions

To illustrate the general theory outlined in the previous subsections we consider here the specific case of a minimally coupled ($\xi = 0$) scalar field in a cosmological spacetime with metric

$$ds^2 = a^2(\eta)(-d\eta^2 + dx^2 + dy^2 + dz^2), \tag{14.34}$$

where $a(\eta)$ is the scale factor expressed in terms of conformal time. For concreteness we take the universe to be matter dominated, so that $a(\eta) = C\eta^2$, where C is a constant. This spacetime is one of the very few for which Green's functions can be explicitly constructed. The calculation presented here was first carried out by Burko, Harte, and Poisson [47]; it can be extended to other cosmologies [48].

To solve Green's equation $\Box G(x,x') = -4\pi\delta_4(x,x')$ we first introduce a reduced Green's function $g(x,x')$ defined by

$$G(x,x') = \frac{g(x,x')}{a(\eta)a(\eta')}. \tag{14.35}$$

Substitution yields

$$\left(-\frac{\partial^2}{\partial\eta^2} + \nabla^2 + \frac{2}{\eta^2}\right)g(x,x') = -4\pi\delta(\eta-\eta')\delta_3(\boldsymbol{x}-\boldsymbol{x'}), \tag{14.36}$$

where $\boldsymbol{x} = (x,y,z)$ is a vector in three-dimensional flat space, and ∇^2 is the Laplacian operator in this space. We next expand $g(x,x')$ in terms of plane-wave solutions to Laplace's equation,

$$g(x,x') = \frac{1}{(2\pi)^3}\int \tilde{g}(\eta,\eta';\boldsymbol{k})\, e^{i\boldsymbol{k}\cdot(\boldsymbol{x}-\boldsymbol{x'})}\, d^3k, \tag{14.37}$$

and we substitute this back into Eq. (14.36). The result, after also Fourier transforming $\delta_3(\boldsymbol{x}-\boldsymbol{x'})$, is an ordinary differential equation for $\tilde{g}(\eta,\eta';\boldsymbol{k})$:

$$\left(\frac{d^2}{d\eta^2} + k^2 - \frac{2}{\eta^2}\right)\tilde{g} = 4\pi\delta(\eta-\eta'), \tag{14.38}$$

where $k^2 = \boldsymbol{k}\cdot\boldsymbol{k}$. To generate the retarded Green's function we set

$$\tilde{g}_+(\eta,\eta';\boldsymbol{k}) = \theta(\eta-\eta')\,\hat{g}(\eta,\eta';k), \tag{14.39}$$

in which we indicate that $\hat{g}$ depends only on the modulus of the vector $\boldsymbol{k}$. To generate the advanced Green's function we would set instead $\tilde{g}_-(\eta,\eta';\boldsymbol{k}) = \theta(\eta'-\eta)\,\hat{g}(\eta,\eta';k)$. The following manipulations will refer specifically to the retarded Green's function; they are easily adapted to the case of the advanced Green's function.

Substitution of Eq. (14.39) into Eq. (14.38) reveals that $\hat{g}$ must satisfy the homogeneous equation

$$\left(\frac{d^2}{d\eta^2} + k^2 - \frac{2}{\eta^2}\right)\hat{g} = 0, \tag{14.40}$$

together with the boundary conditions

$$\hat{g}(\eta=\eta';k) = 0, \qquad \frac{d\hat{g}}{d\eta}(\eta=\eta';k) = 4\pi. \tag{14.41}$$

Inserting Eq. (14.39) into Eq. (14.37) and integrating over the angular variables associated with the vector $\boldsymbol{k}$ yields

$$g_+(x,x') = \frac{\theta(\Delta\eta)}{2\pi^2 R}\int_0^\infty \hat{g}(\eta,\eta';k)\,k\sin(kR)\,dk, \tag{14.42}$$

where $\Delta\eta := \eta-\eta'$ and $R := |\boldsymbol{x}-\boldsymbol{x'}|$.

Equation (14.40) has $\cos(k\Delta\eta) - (k\eta)^{-1}\sin(k\Delta\eta)$ and $\sin(k\Delta\eta) + (k\eta)^{-1}\cos(k\Delta\eta)$ as linearly independent solutions, and $\hat{g}(\eta,\eta';k)$ must be given by a linear superposition. The coefficients can be functions of η', and after imposing Eqs. (14.41) we find that the appropriate combination is

$$\hat{g}(\eta,\eta';k) = \frac{4\pi}{k}\left[\left(1+\frac{1}{k^2\eta\eta'}\right)\sin(k\Delta\eta) - \frac{\Delta\eta}{k\eta\eta'}\cos(k\Delta\eta)\right]. \tag{14.43}$$

Substituting this into Eq. (14.42) and using the identity $(2/\pi)\int_0^\infty \sin(\omega x)\sin(\omega x')\,d\omega = \delta(x-x') - \delta(x+x')$ yields

$$g_+(x,x') = \frac{\delta(\Delta\eta - R)}{R} + \frac{\theta(\Delta\eta)}{\eta\eta'}\,\frac{2}{\pi}\int_0^\infty \frac{1}{k}\,\sin(k\Delta\eta)\cos(kR)\,dk$$

after integration by parts. The integral evaluates to $\theta(\Delta\eta - R)$.

We have arrived at

$$g_+(x,x') = \frac{\delta(\eta-\eta'-|\boldsymbol{x}-\boldsymbol{x'}|)}{|\boldsymbol{x}-\boldsymbol{x'}|} + \frac{\theta(\eta-\eta'-|\boldsymbol{x}-\boldsymbol{x'}|)}{\eta\eta'} \tag{14.44}$$

for our final expression for the retarded Green's function. The advanced Green's function is given instead by

$$g_-(x,x') = \frac{\delta(\eta-\eta'+|\boldsymbol{x}-\boldsymbol{x'}|)}{|\boldsymbol{x}-\boldsymbol{x'}|} + \frac{\theta(-\eta+\eta'-|\boldsymbol{x}-\boldsymbol{x'}|)}{\eta\eta'}. \tag{14.45}$$

The distributions $g_\pm(x,x')$ are solutions to the reduced Green's equation of Eq. (14.36). The actual Green's functions are obtained by substituting Eqs. (14.44) and (14.45) into Eq. (14.35). We note that the support of the retarded Green's function is given by $\eta-\eta' \geq |\boldsymbol{x}-\boldsymbol{x'}|$, while the support of the advanced Green's function is given by $\eta-\eta' \leq -|\boldsymbol{x}-\boldsymbol{x'}|$.

It may be verified that the symmetric two-point function

$$h(x,x') = \frac{1}{\eta\eta'} \tag{14.46}$$

satisfies all of the properties H1–H4 listed in Sec. 14.5; it may thus be used to define singular and regular Green's functions. According to Eq. (14.30) the singular Green's function is given by

$$\begin{aligned} g_{\rm S}(x,x') \quad &= \quad \frac{1}{2|\boldsymbol{x}-\boldsymbol{x'}|}\Big[\delta(\eta-\eta'-|\boldsymbol{x}-\boldsymbol{x'}|)+\delta(\eta-\eta'+|\boldsymbol{x}-\boldsymbol{x'}|)\Big] \\ &\quad + \frac{1}{2\eta\eta'}\Big[\theta(\eta-\eta'-|\boldsymbol{x}-\boldsymbol{x'}|)-\theta(\eta-\eta'+|\boldsymbol{x}-\boldsymbol{x'}|)\Big] \end{aligned} \tag{14.47}$$

and its support is limited to the interval $-|\boldsymbol{x}-\boldsymbol{x'}| \leq \eta-\eta' \leq |\boldsymbol{x}-\boldsymbol{x'}|$. According to Eq. (14.22) the regular two-point function is given by

$$\begin{aligned} g_{\rm R}(x,x') \quad &= \quad \frac{1}{2|\boldsymbol{x}-\boldsymbol{x'}|}\Big[\delta(\eta-\eta'-|\boldsymbol{x}-\boldsymbol{x'}|)-\delta(\eta-\eta'+|\boldsymbol{x}-\boldsymbol{x'}|)\Big] \\ &\quad + \frac{1}{2\eta\eta'}\Big[\theta(\eta-\eta'-|\boldsymbol{x}-\boldsymbol{x'}|)+\theta(\eta-\eta'+|\boldsymbol{x}-\boldsymbol{x'}|)\Big]; \end{aligned} \tag{14.48}$$

its support is given by $\eta-\eta' \geq -|\boldsymbol{x}-\boldsymbol{x'}|$ and for $\eta-\eta' \geq |\boldsymbol{x}-\boldsymbol{x'}|$ the regular two-point function agrees with the retarded Green's function.

As a final observation we note that for this cosmological spacetime, the normal convex neighbourhood of any point x consists of the whole spacetime manifold (which excludes the cosmological singularity at $a = 0$). The Hadamard construction of the Green's functions is therefore valid globally, a fact that is immediately revealed by Eqs. (14.44) and (14.45).

15 Electromagnetic Green's functions

15.1 Equations of electromagnetism

The electromagnetic field tensor $F_{\alpha\beta} = \nabla_\alpha A_\beta - \nabla_\beta A_\alpha$ is expressed in terms of a vector potential A_α. In the Lorenz gauge $\nabla_\alpha A^\alpha = 0$, the vector potential satisfies the wave equation

$$\Box A^\alpha - R^\alpha_{\ \beta} A^\beta = -4\pi j^\alpha, \tag{15.1}$$

where $\Box = g^{\alpha\beta}\nabla_\alpha\nabla_\beta$ is the wave operator, $R^\alpha_{\ \beta}$ the Ricci tensor, and j^α a prescribed current density. The wave equation enforces the condition $\nabla_\alpha j^\alpha = 0$, which expresses charge conservation.

The solution to the wave equation is written as

$$A^\alpha(x) = \int G^\alpha_{\ \beta'}(x,x') j^{\beta'}(x') \sqrt{-g'}\, d^4x', \tag{15.2}$$

in terms of a Green's function $G^\alpha_{\ \beta'}(x,x')$ that satisfies

$$\Box G^\alpha_{\ \beta'}(x,x') - R^\alpha_{\ \beta}(x) G^\beta_{\ \beta'}(x,x') = -4\pi g^\alpha_{\ \beta'}(x,x')\delta_4(x,x'), \tag{15.3}$$

where $g^{\alpha}{}_{\beta'}(x,x')$ is a parallel propagator and $\delta_4(x,x')$ an invariant Dirac distribution. The parallel propagator is inserted on the right-hand side of Eq. (15.3) to keep the index structure of the equation consistent from side to side; because $g^{\alpha}{}_{\beta'}(x,x')\delta_4(x,x')$ is distributionally equal to $[g^{\alpha}{}_{\beta'}]\delta_4(x,x') = \delta^{\alpha'}{}_{\beta'}\delta_4(x,x')$, it could have been replaced by either $\delta^{\alpha'}{}_{\beta'}$ or $\delta^{\alpha}{}_{\beta}$. It is easy to check that by virtue of Eq. (15.3), the vector potential of Eq. (15.2) satisfies the wave equation of Eq. (15.1).

We will assume that the retarded Green's function $G_{+}{}^{\alpha}{}_{\beta'}(x,x')$, which is nonzero if x is in the causal future of x', and the advanced Green's function $G_{-}{}^{\alpha}{}_{\beta'}(x,x')$, which is nonzero if x is in the causal past of x', exist as distributions and can be defined globally in the entire spacetime.

15.2 Hadamard construction of the Green's functions

Assuming throughout this subsection that x is in the normal convex neighbourhood of x', we make the ansatz

$$G_{\pm}{}^{\alpha}{}_{\beta'}(x,x') = U^{\alpha}{}_{\beta'}(x,x')\delta_{\pm}(\sigma) + V^{\alpha}{}_{\beta'}(x,x')\theta_{\pm}(-\sigma), \tag{15.4}$$

where $\theta_{\pm}(-\sigma)$, $\delta_{\pm}(\sigma)$ are the light-cone distributions introduced in Sec. 13.2, and where $U^{\alpha}{}_{\beta'}(x,x')$, $V^{\alpha}{}_{\beta'}(x,x')$ are smooth bitensors.

To conveniently manipulate the Green's functions we shift σ by a small positive quantity ϵ. The Green's functions are then recovered by the taking the limit of

$$G^{\epsilon}_{\pm}{}^{\alpha}{}_{\beta'}(x,x') := U^{\alpha}{}_{\beta'}(x,x')\delta_{\pm}(\sigma+\epsilon) + V^{\alpha}{}_{\beta'}(x,x')\theta_{\pm}(-\sigma-\epsilon)$$

as $\epsilon \to 0^+$. When we substitute this into the left-hand side of Eq. (15.3) and then take the limit, we obtain

$$\begin{aligned}
\Box G_{\pm}{}^{\alpha}{}_{\beta'} - R^{\alpha}{}_{\beta} G_{\pm}{}^{\beta}{}_{\beta'} \quad = \quad & -4\pi\delta_4(x,x')U^{\alpha}{}_{\beta'} + \delta'_{\pm}(\sigma)\Big\{2U^{\alpha}{}_{\beta';\gamma}\sigma^{\gamma} + (\sigma^{\gamma}{}_{\gamma} - 4)U^{\alpha}{}_{\beta'}\Big\} \\
& + \delta_{\pm}(\sigma)\Big\{-2V^{\alpha}{}_{\beta';\gamma}\sigma^{\gamma} + (2-\sigma^{\gamma}{}_{\gamma})V^{\alpha}{}_{\beta'} + \Box U^{\alpha}{}_{\beta'} - R^{\alpha}{}_{\beta}U^{\beta}{}_{\beta'}\Big\} \\
& + \theta_{\pm}(-\sigma)\Big\{\Box V^{\alpha}{}_{\beta'} - R^{\alpha}{}_{\beta}V^{\beta}{}_{\beta'}\Big\}
\end{aligned}$$

after a routine computation similar to the one presented at the beginning of Sec. 14.2. Comparison with Eq. (15.3) returns: (i) the equations

$$\big[U^{\alpha}{}_{\beta'}\big] = \big[g^{\alpha}{}_{\beta'}\big] = \delta^{\alpha'}{}_{\beta'} \tag{15.5}$$

and

$$2U^{\alpha}{}_{\beta';\gamma}\sigma^{\gamma} + (\sigma^{\gamma}{}_{\gamma} - 4)U^{\alpha}{}_{\beta'} = 0 \tag{15.6}$$

that determine $U^{\alpha}{}_{\beta'}(x,x')$; (ii) the equation

$$\check{V}^{\alpha}{}_{\beta';\gamma}\sigma^{\gamma} + \frac{1}{2}(\sigma^{\gamma}{}_{\gamma} - 2)\check{V}^{\alpha}{}_{\beta'} = \frac{1}{2}\big(\Box U^{\alpha}{}_{\beta'} - R^{\alpha}{}_{\beta}U^{\beta}{}_{\beta'}\big)\Big|_{\sigma=0} \tag{15.7}$$

that determines $\check{V}^{\alpha}{}_{\beta'}(x,x')$, the restriction of $V^{\alpha}{}_{\beta'}(x,x')$ on the light cone $\sigma(x,x') = 0$; and (iii) the wave equation

$$\Box V^{\alpha}{}_{\beta'} - R^{\alpha}{}_{\beta}V^{\beta}{}_{\beta'} = 0 \tag{15.8}$$

that determines $V^{\alpha}{}_{\beta'}(x,x')$ inside the light cone.

Equation (15.6) can be integrated along the unique geodesic β that links x' to x. The initial conditions are provided by Eq. (15.5), and if we set $U^{\alpha}{}_{\beta'}(x,x') = g^{\alpha}{}_{\beta'}(x,x')U(x,x')$, we find that these equations reduce to Eqs. (14.7) and (14.6), respectively. According to Eq. (14.8), then, we have

$$U^{\alpha}{}_{\beta'}(x,x') = g^{\alpha}{}_{\beta'}(x,x')\Delta^{1/2}(x,x'), \tag{15.9}$$

which reduces to

$$U^{\alpha}{}_{\beta'} = g^{\alpha}{}_{\beta'}\Big(1 + \frac{1}{12}R_{\gamma'\delta'}\sigma^{\gamma'}\sigma^{\delta'} + O(\lambda^3)\Big) \tag{15.10}$$

near coincidence, with λ denoting the affine-parameter distance between x' and x. Differentiation of this relation gives

$$U^{\alpha}{}_{\beta';\gamma} = \frac{1}{2}g^{\gamma'}{}_{\gamma}\Big(g^{\alpha}{}_{\alpha'}R^{\alpha'}{}_{\beta'\gamma'\delta'} - \frac{1}{3}g^{\alpha}{}_{\beta'}R_{\gamma'\delta'}\Big)\sigma^{\delta'} + O(\lambda^2), \tag{15.11}$$

$$U^{\alpha}{}_{\beta';\gamma'} = \frac{1}{2}\Big(g^{\alpha}{}_{\alpha'}R^{\alpha'}{}_{\beta'\gamma'\delta'} + \frac{1}{3}g^{\alpha}{}_{\beta'}R_{\gamma'\delta'}\Big)\sigma^{\delta'} + O(\lambda^2), \tag{15.12}$$

and eventually,

$$\big[\Box U^{\alpha}{}_{\beta'}\big] = \frac{1}{6}\delta^{\alpha'}{}_{\beta'}R(x'). \tag{15.13}$$

Similarly, Eq. (15.7) can be integrated along each null geodesic that generates the null cone $\sigma(x,x') = 0$. The initial values are obtained by taking the coincidence limit of this equation, using Eqs. (15.5), (15.13), and the additional relation $[\sigma^{\gamma}{}_{\gamma}] = 4$. We arrive at

$$\big[V^{\alpha}{}_{\beta'}\big] = -\frac{1}{2}\Big(R^{\alpha'}{}_{\beta'} - \frac{1}{6}\delta^{\alpha'}{}_{\beta'}R'\Big). \tag{15.14}$$

With the characteristic data obtained by integrating Eq. (15.7), the wave equation of Eq. (15.8) admits a unique solution.

To summarize, the retarded and advanced electromagnetic Green's functions are given by Eq. (15.4) with $U^{\alpha}{}_{\beta'}(x,x')$ given by Eq. (15.9) and $V^{\alpha}{}_{\beta'}(x,x')$ determined by Eq. (15.8) and the characteristic data constructed with Eqs. (15.7) and (15.14). It should be emphasized that the construction provided in this subsection is restricted to $\mathcal{N}(x')$, the normal convex neighbourhood of the reference point x'.

15.3 Reciprocity and Kirchhoff representation

Like their scalar counterparts, the (globally defined) electromagnetic Green's functions satisfy a reciprocity relation, the statement of which is

$$G^{-}_{\beta'\alpha}(x',x) = G^{+}_{\alpha\beta'}(x,x'). \tag{15.15}$$

The derivation of Eq. (15.15) is virtually identical to what was presented in Sec. 14.3, and we shall not present the details. It suffices to mention that it is based on the identities

$$G^{+}_{\alpha\beta'}(x,x')\Big(\Box G_{-}{}^{\alpha}{}_{\gamma''}(x,x'') - R^{\alpha}{}_{\gamma}G_{-}{}^{\gamma}{}_{\gamma''}(x,x'')\Big) = -4\pi G^{+}_{\alpha\beta'}(x,x')g^{\alpha}{}_{\gamma''}(x,x'')\delta_4(x,x'')$$

and

$$G^{-}_{\alpha\gamma''}(x,x'')\Big(\Box G_{+}{}^{\alpha}{}_{\beta'}(x,x') - R^{\alpha}{}_{\gamma}G_{+}{}^{\gamma}{}_{\beta'}(x,x')\Big) = -4\pi G^{-}_{\alpha\gamma''}(x,x'')g^{\alpha}{}_{\beta'}(x,x')\delta_4(x,x').$$

A direct consequence of the reciprocity relation is

$$V_{\beta'\alpha}(x',x) = V_{\alpha\beta'}(x,x'), \tag{15.16}$$

the statement that the bitensor $V_{\alpha\beta'}(x,x')$ is symmetric in its indices and arguments.

The Kirchhoff representation for the electromagnetic vector potential is formulated as follows. Suppose that $A^{\alpha}(x)$ satisfies the *homogeneous* version of Eq. (15.1) and that initial values $A^{\alpha'}(x')$, $n^{\beta'}\nabla_{\beta'}A^{\alpha'}(x')$ are specified on a spacelike hypersurface Σ. Then the value of the potential at a point x in the future of Σ is given by

$$A^{\alpha}(x) = -\frac{1}{4\pi}\int_{\Sigma}\Big(G_{+}{}^{\alpha}{}_{\beta'}(x,x')\nabla^{\gamma'}A^{\beta'}(x') - A^{\beta'}(x')\nabla^{\gamma'}G_{+}{}^{\alpha}{}_{\beta'}(x,x')\Big)\, d\Sigma_{\gamma'}, \tag{15.17}$$

where $d\Sigma_{\gamma'} = -n_{\gamma'}dV$ is a surface element on Σ; $n_{\gamma'}$ is the future-directed unit normal and dV is the invariant volume element on the hypersurface. The derivation of Eq. (15.17) is virtually identical to what was presented in Sec. 14.4.

15.4 Relation with scalar Green's functions

In a spacetime that satisfies the Einstein field equations in vacuum, so that $R_{\alpha\beta} = 0$ everywhere in the spacetime, the (retarded and advanced) electromagnetic Green's functions satisfy the identities [4]

$$G_{\pm\ \beta';\alpha}^{\ \alpha} = -G_{\pm;\beta'}, \tag{15.18}$$

where $G_{\pm}$ are the corresponding scalar Green's functions.

To prove this we differentiate Eq. (15.3) covariantly with respect to x^{α} and use Eq. (13.3) to express the right-hand side as $+4\pi\partial_{\beta'}\delta_4(x,x')$. After repeated use of Ricci's identity to permute the ordering of the covariant derivatives on the left-hand side, we arrive at the equation

$$\Box\big(-G^{\alpha}_{\ \beta';\alpha}\big) = -4\pi\partial_{\beta'}\delta_4(x,x'); \tag{15.19}$$

all terms involving the Riemann tensor disappear by virtue of the fact that the spacetime is Ricci-flat. Because Eq. (15.19) is also the differential equation satisfied by $G_{;\beta'}$, and because the solutions are chosen to satisfy the same boundary conditions, we have established the validity of Eq. (15.18).

15.5 Singular and regular Green's functions

We shall now construct singular and regular Green's functions for the electromagnetic field. The treatment here parallels closely what was presented in Sec. 14.5, and the reader is referred to that section for a more complete discussion.

We begin by introducing the bitensor $H^{\alpha}_{\ \beta'}(x,x')$ with properties

H1: $H^{\alpha}_{\ \beta'}(x,x')$ satisfies the homogeneous wave equation,

$$\Box H^{\alpha}_{\ \beta'}(x,x') - R^{\alpha}_{\ \beta}(x)H^{\beta}_{\ \beta'}(x,x') = 0; \tag{15.20}$$

H2: $H^{\alpha}_{\ \beta'}(x,x')$ is symmetric in its indices and arguments,

$$H_{\beta'\alpha}(x',x) = H_{\alpha\beta'}(x,x'); \tag{15.21}$$

H3: $H^{\alpha}_{\ \beta'}(x,x')$ agrees with the retarded Green's function if x is in the chronological future of x',

$$H^{\alpha}_{\ \beta'}(x,x') = G_{+\ \beta'}^{\ \alpha}(x,x') \qquad \text{when } x \in I^{+}(x'); \tag{15.22}$$

H4: $H^{\alpha}_{\ \beta'}(x,x')$ agrees with the advanced Green's function if x is in the chronological past of x',

$$H^{\alpha}_{\ \beta'}(x,x') = G_{-\ \beta'}^{\ \alpha}(x,x') \qquad \text{when } x \in I^{-}(x'). \tag{15.23}$$

It is easy to prove that property H4 follows from H2, H3, and the reciprocity relation (15.15) satisfied by the retarded and advanced Green's functions. That such a bitensor exists can be argued along the same lines as those presented in Sec. 14.5.

Equipped with the bitensor $H^{\alpha}_{\ \beta'}(x,x')$ we define the singular Green's function to be

$$G_{\mathrm{S}\ \beta'}^{\ \alpha}(x,x') = \frac{1}{2}\Big[G_{+\ \beta'}^{\ \alpha}(x,x') + G_{-\ \beta'}^{\ \alpha}(x,x') - H^{\alpha}_{\ \beta'}(x,x')\Big]. \tag{15.24}$$

This comes with the properties

S1: $G_{\mathrm{S}\ \beta'}^{\ \alpha}(x,x')$ satisfies the inhomogeneous wave equation,

$$\Box G_{\mathrm{S}\ \beta'}^{\ \alpha}(x,x') - R^{\alpha}_{\ \beta}(x)G_{\mathrm{S}\ \beta'}^{\ \beta}(x,x') = -4\pi g^{\alpha}_{\ \beta'}(x,x')\delta_4(x,x'); \tag{15.25}$$

S2: $G_{\mathrm{S}\ \beta'}^{\ \alpha}(x,x')$ is symmetric in its indices and arguments,

$$G^{\mathrm{S}}_{\beta'\alpha}(x',x) = G^{\mathrm{S}}_{\alpha\beta'}(x,x'); \tag{15.26}$$

S3: $G_{\mathrm{S}}{}^{\alpha}{}_{\beta'}(x,x')$ vanishes if x is in the chronological past or future of x',

$$G_{\mathrm{S}}{}^{\alpha}{}_{\beta'}(x,x') = 0 \qquad \text{when } x \in I^{\pm}(x'). \tag{15.27}$$

These can be established as consequences of H1–H4 and the properties of the retarded and advanced Green's functions.

The regular two-point function is then defined by

$$G_{\mathrm{R}}{}^{\alpha}{}_{\beta'}(x,x') = G_{+}{}^{\alpha}{}_{\beta'}(x,x') - G_{\mathrm{S}}{}^{\alpha}{}_{\beta'}(x,x'), \tag{15.28}$$

and it comes with the properties

R1: $G_{\mathrm{R}}{}^{\alpha}{}_{\beta'}(x,x')$ satisfies the homogeneous wave equation,

$$\Box G_{\mathrm{R}}{}^{\alpha}{}_{\beta'}(x,x') - R^{\alpha}{}_{\beta}(x) G_{\mathrm{R}}{}^{\beta}{}_{\beta'}(x,x') = 0; \tag{15.29}$$

R2: $G_{\mathrm{R}}{}^{\alpha}{}_{\beta'}(x,x')$ agrees with the retarded Green's function if x is in the chronological future of x',

$$G_{\mathrm{R}}{}^{\alpha}{}_{\beta'}(x,x') = G_{+}{}^{\alpha}{}_{\beta'}(x,x') \qquad \text{when } x \in I^{+}(x'); \tag{15.30}$$

R3: $G_{\mathrm{R}}{}^{\alpha}{}_{\beta'}(x,x')$ vanishes if x is in the chronological past of x',

$$G_{\mathrm{R}}{}^{\alpha}{}_{\beta'}(x,x') = 0 \qquad \text{when } x \in I^{-}(x'). \tag{15.31}$$

Those follow immediately from S1–S3 and the properties of the retarded Green's function.

When x is restricted to the normal convex neighbourhood of x', we have the explicit relations

$$\begin{aligned}
H^{\alpha}{}_{\beta'}(x,x') &= V^{\alpha}{}_{\beta'}(x,x'), && (15.32)\\
G_{\mathrm{S}}{}^{\alpha}{}_{\beta'}(x,x') &= \frac{1}{2}U^{\alpha}{}_{\beta'}(x,x')\delta(\sigma) - \frac{1}{2}V^{\alpha}{}_{\beta'}(x,x')\theta(\sigma), && (15.33)\\
G_{\mathrm{R}}{}^{\alpha}{}_{\beta'}(x,x') &= \frac{1}{2}U^{\alpha}{}_{\beta'}(x,x')\Big[\delta_{+}(\sigma) - \delta_{-}(\sigma)\Big] + V^{\alpha}{}_{\beta'}(x,x')\Big[\theta_{+}(-\sigma) + \frac{1}{2}\theta(\sigma)\Big]. && (15.34)
\end{aligned}$$

From these we see clearly that the singular Green's function does not distinguish between past and future (property S2), and that its support excludes $I^{\pm}(x')$ (property S3). We see also that the regular two-point function coincides with $G_{+}{}^{\alpha}{}_{\beta'}(x,x')$ in $I^{+}(x')$ (property R2), and that its support does not include $I^{-}(x')$ (property R3).

16 Gravitational Green's functions

16.1 Equations of linearized gravity

We are given a background spacetime for which the metric $g_{\alpha\beta}$ satisfies the Einstein field equations *in vacuum*. We then perturb the metric from $g_{\alpha\beta}$ to

$$\mathsf{g}_{\alpha\beta} = g_{\alpha\beta} + h_{\alpha\beta}. \tag{16.1}$$

The metric perturbation $h_{\alpha\beta}$ is assumed to be small, and when working out the Einstein field equations to be satisfied by the new metric $\mathsf{g}_{\alpha\beta}$, we work consistently to first order in $h_{\alpha\beta}$. To simplify the expressions we use the trace-reversed potentials $\gamma_{\alpha\beta}$ defined by

$$\gamma_{\alpha\beta} = h_{\alpha\beta} - \frac{1}{2}\big(g^{\gamma\delta}h_{\gamma\delta}\big)g_{\alpha\beta}, \tag{16.2}$$

and we impose the Lorenz gauge condition,

$$\gamma^{\alpha\beta}{}_{;\beta} = 0. \tag{16.3}$$

In this equation, and in all others below, indices are raised and lowered with the background metric $g_{\alpha\beta}$. Similarly, the connection involved in Eq. (16.3), and in all other equations below, is the one that is compatible with the background metric. If $T^{\alpha\beta}$ is the perturbing energy-momentum tensor, then by virtue of the linearized Einstein field equations the perturbation field obeys the wave equation

$$\Box\gamma^{\alpha\beta} + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}\gamma^{\gamma\delta} = -16\pi T^{\alpha\beta}, \tag{16.4}$$

in which $\Box = g^{\alpha\beta}\nabla_\alpha\nabla_\beta$ is the wave operator and $R_{\gamma\alpha\delta\beta}$ the Riemann tensor. In first-order perturbation theory, the energy-momentum tensor must be conserved in the background spacetime: $T^{\alpha\beta}_{\ \ \ ;\beta} = 0$.

The solution to the wave equation is written as

$$\gamma^{\alpha\beta}(x) = 4\int G^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')T^{\gamma'\delta'}(x')\sqrt{-g'}\,d^4x', \tag{16.5}$$

in terms of a Green's function $G^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ that satisfies [140]

$$\Box G^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x') + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}(x)G^{\gamma\delta}_{\ \ \gamma'\delta'}(x,x') = -4\pi g^{(\alpha}_{\ \ \gamma'}(x,x')g^{\beta)}_{\ \ \delta'}(x,x')\delta_4(x,x'), \tag{16.6}$$

where $g^{\alpha}_{\ \gamma'}(x,x')$ is a parallel propagator and $\delta_4(x,x')$ an invariant Dirac functional. The parallel propagators are inserted on the right-hand side of Eq. (16.6) to keep the index structure of the equation consistent from side to side; in particular, both sides of the equation are symmetric in α and β, and in γ' and δ'. It is easy to check that by virtue of Eq. (16.6), the perturbation field of Eq. (16.5) satisfies the wave equation of Eq. (16.4). Once $\gamma_{\alpha\beta}$ is known, the metric perturbation can be reconstructed from the relation $h_{\alpha\beta} = \gamma_{\alpha\beta} - \frac{1}{2}(g^{\gamma\delta}\gamma_{\gamma\delta})g_{\alpha\beta}$.

We will assume that the retarded Green's function $G_{+\ \gamma'\delta'}^{\ \alpha\beta}(x,x')$, which is nonzero if x is in the causal future of x', and the advanced Green's function $G_{-\ \gamma'\delta'}^{\ \alpha\beta}(x,x')$, which is nonzero if x is in the causal past of x', exist as distributions and can be defined globally in the entire background spacetime.

16.2 Hadamard construction of the Green's functions

Assuming throughout this subsection that x is in the normal convex neighbourhood of x', we make the ansatz

$$G_{\pm\ \gamma'\delta'}^{\ \alpha\beta}(x,x') = U^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')\delta_\pm(\sigma) + V^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')\theta_\pm(-\sigma), \tag{16.7}$$

where $\theta_\pm(-\sigma)$, $\delta_\pm(\sigma)$ are the light-cone distributions introduced in Sec. 13.2, and where $U^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$, $V^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ are smooth bitensors.

To conveniently manipulate the Green's functions we shift σ by a small positive quantity ϵ. The Green's functions are then recovered by the taking the limit of

$$G^{\epsilon\ \alpha\beta}_{\pm\ \gamma'\delta'}(x,x') = U^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')\delta_\pm(\sigma+\epsilon) + V^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')\theta_\pm(-\sigma-\epsilon)$$

as $\epsilon \to 0^+$. When we substitute this into the left-hand side of Eq. (16.6) and then take the limit, we obtain

$$\begin{aligned}
\Box G_{\pm\ \gamma'\delta'}^{\ \alpha\beta} + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}G_{\pm\ \gamma'\delta'}^{\ \gamma\delta} \quad = \quad & -4\pi\delta_4(x,x')U^{\alpha\beta}_{\ \ \gamma'\delta'} + \delta'_\pm(\sigma)\Big\{2U^{\alpha\beta}_{\ \ \gamma'\delta';\gamma}\sigma^\gamma + (\sigma^\gamma_{\ \gamma} - 4)U^{\alpha\beta}_{\ \ \gamma'\delta'}\Big\} \\
& + \delta_\pm(\sigma)\Big\{-2V^{\alpha\beta}_{\ \ \gamma'\delta';\gamma}\sigma^\gamma + (2-\sigma^\gamma_{\ \gamma})V^{\alpha\beta}_{\ \ \gamma'\delta'} + \Box U^{\alpha\beta}_{\ \ \gamma'\delta'} + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}U^{\gamma\delta}_{\ \ \gamma'\delta'}\Big\} \\
& + \theta_\pm(-\sigma)\Big\{\Box V^{\alpha\beta}_{\ \ \gamma'\delta'} + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}V^{\gamma\delta}_{\ \ \gamma'\delta'}\Big\}
\end{aligned}$$

after a routine computation similar to the one presented at the beginning of Sec. 14.2. Comparison with Eq. (16.6) returns: (i) the equations

$$\big[U^{\alpha\beta}_{\ \ \gamma'\delta'}\big] = \Big[g^{(\alpha}_{\ \ \gamma'}g^{\beta)}_{\ \ \delta'}\Big] = \delta^{(\alpha'}_{\ \ \gamma'}\delta^{\beta')}_{\ \ \delta'} \tag{16.8}$$

and

$$2U^{\alpha\beta}_{\ \ \gamma'\delta';\gamma}\sigma^\gamma + (\sigma^\gamma_{\ \gamma} - 4)U^{\alpha\beta}_{\ \ \gamma'\delta'} = 0 \tag{16.9}$$

that determine $U^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$; (ii) the equation

$$\check{V}^{\alpha\beta}_{\ \ \gamma'\delta';\gamma}\sigma^{\gamma} + \frac{1}{2}(\sigma^{\gamma}_{\ \gamma} - 2)\check{V}^{\alpha\beta}_{\ \ \gamma'\delta'} = \frac{1}{2}\bigl(\Box U^{\alpha\beta}_{\ \ \gamma'\delta'} + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}U^{\gamma\delta}_{\ \ \gamma'\delta'}\bigr)\Big|_{\sigma=0} \tag{16.10}$$

that determine $\check{V}^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$, the restriction of $V^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ on the light cone $\sigma(x,x')=0$; and (iii) the wave equation

$$\Box V^{\alpha\beta}_{\ \ \gamma'\delta'} + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}V^{\gamma\delta}_{\ \ \gamma'\delta'} = 0 \tag{16.11}$$

that determines $V^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ inside the light cone.

Equation (16.9) can be integrated along the unique geodesic β that links x' to x. The initial conditions are provided by Eq. (16.8), and if we set $U^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x') = g^{(\alpha}_{\ \ \gamma'}g^{\beta)}_{\ \ \delta'}U(x,x')$, we find that these equations reduce to Eqs. (14.7) and (14.6), respectively. According to Eq. (14.8), then, we have

$$U^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x') = g^{(\alpha}_{\ \ \gamma'}(x,x')g^{\beta)}_{\ \ \delta'}(x,x')\Delta^{1/2}(x,x'), \tag{16.12}$$

which reduces to

$$U^{\alpha\beta}_{\ \ \gamma'\delta'} = g^{(\alpha}_{\ \ \gamma'}g^{\beta)}_{\ \ \delta'}\Bigl(1 + O(\lambda^3)\Bigr) \tag{16.13}$$

near coincidence, with λ denoting the affine-parameter distance between x' and x; there is no term of order λ^2 because by assumption, the background Ricci tensor vanishes at x' (as it does in the entire spacetime). Differentiation of this relation gives

$$U^{\alpha\beta}_{\ \ \gamma'\delta';\epsilon} = \frac{1}{2}g^{(\alpha}_{\ \ \alpha'}g^{\beta)}_{\ \ \beta'}g^{\epsilon'}_{\ \ \epsilon}\Bigl(R^{\alpha'}_{\ \ \gamma'\epsilon'\iota'}\delta^{\beta'}_{\ \ \delta'} + R^{\alpha'}_{\ \ \delta'\epsilon'\iota'}\delta^{\beta'}_{\ \ \gamma'}\Bigr)\sigma^{\iota'} + O(\lambda^2), \tag{16.14}$$

$$U^{\alpha\beta}_{\ \ \gamma'\delta';\epsilon'} = \frac{1}{2}g^{(\alpha}_{\ \ \alpha'}g^{\beta)}_{\ \ \beta'}\Bigl(R^{\alpha'}_{\ \ \gamma'\epsilon'\iota'}\delta^{\beta'}_{\ \ \delta'} + R^{\alpha'}_{\ \ \delta'\epsilon'\iota'}\delta^{\beta'}_{\ \ \gamma'}\Bigr)\sigma^{\iota'} + O(\lambda^2), \tag{16.15}$$

and eventually,

$$\bigl[\Box U^{\alpha\beta}_{\ \ \gamma'\delta'}\bigr] = 0; \tag{16.16}$$

this last result follows from the fact that $[U^{\alpha\beta}_{\ \ \gamma'\delta';\epsilon\iota}]$ is antisymmetric in the last pair of indices.

Similarly, Eq. (16.10) can be integrated along each null geodesic that generates the null cone $\sigma(x,x')=0$. The initial values are obtained by taking the coincidence limit of this equation, using Eqs. (16.8), (16.16), and the additional relation $[\sigma^{\gamma}_{\ \gamma}] = 4$. We arrive at

$$\bigl[V^{\alpha\beta}_{\ \ \gamma'\delta'}\bigr] = \frac{1}{2}\Bigl(R^{\alpha'\ \ \beta'}_{\ \ \gamma'\ \ \delta'} + R^{\beta'\ \ \alpha'}_{\ \ \gamma'\ \ \delta'}\Bigr). \tag{16.17}$$

With the characteristic data obtained by integrating Eq. (16.10), the wave equation of Eq. (16.11) admits a unique solution.

To summarize, the retarded and advanced gravitational Green's functions are given by Eq. (16.7) with $U^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ given by Eq. (16.12) and $V^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ determined by Eq. (16.11) and the characteristic data constructed with Eqs. (16.10) and (16.17). It should be emphasized that the construction provided in this subsection is restricted to $\mathcal{N}(x')$, the normal convex neighbourhood of the reference point x'.

16.3 Reciprocity and Kirchhoff representation

The (globally defined) gravitational Green's functions satisfy the reciprocity relation

$$G^{-}_{\gamma'\delta'\alpha\beta}(x',x) = G^{+}_{\alpha\beta\gamma'\delta'}(x,x'). \tag{16.18}$$

The derivation of this result is virtually identical to what was presented in Secs. 14.3 and 15.3. A direct consequence of the reciprocity relation is the statement

$$V_{\gamma'\delta'\alpha\beta}(x',x) = V_{\alpha\beta\gamma'\delta'}(x,x'). \tag{16.19}$$

The Kirchhoff representation for the trace-reversed gravitational perturbation $\gamma_{\alpha\beta}$ is formulated as follows. Suppose that $\gamma^{\alpha\beta}(x)$ satisfies the homogeneous version of Eq. (16.4) and that initial values $\gamma^{\alpha'\beta'}(x')$,

$n^{\gamma'}\nabla_{\gamma'}\gamma^{\alpha'\beta'}(x')$ are specified on a spacelike hypersurface Σ. Then the value of the perturbation field at a point x in the future of Σ is given by

$$\gamma^{\alpha\beta}(x) = -\frac{1}{4\pi}\int_\Sigma \Big(G_{+\ \gamma'\delta'}^{\ \alpha\beta}(x,x')\nabla^{\epsilon'}\gamma^{\gamma'\delta'}(x') - \gamma^{\gamma'\delta'}(x')\nabla^{\epsilon'}G_{+\ \gamma'\delta'}^{\ \alpha\beta}(x,x')\Big)\, d\Sigma_{\epsilon'}, \tag{16.20}$$

where $d\Sigma_{\epsilon'} = -n_{\epsilon'}dV$ is a surface element on Σ; $n_{\epsilon'}$ is the future-directed unit normal and dV is the invariant volume element on the hypersurface. The derivation of Eq. (16.20) is virtually identical to what was presented in Secs. 14.4 and 15.3.

16.4 Relation with electromagnetic and scalar Green's functions

In a spacetime that satisfies the Einstein field equations in vacuum, so that $R_{\alpha\beta} = 0$ everywhere in the spacetime, the (retarded and advanced) gravitational Green's functions satisfy the identities [17]

$$G_{\pm\ \gamma'\delta';\beta}^{\ \alpha\beta} = -G_{\pm(\gamma';\delta')}^{\ \alpha} \tag{16.21}$$

and

$$g^{\gamma'\delta'}G_{\pm\ \gamma'\delta'}^{\ \alpha\beta} = g^{\alpha\beta}G_\pm, \tag{16.22}$$

where $G_{\pm\beta'}^{\ \alpha}$ are the corresponding electromagnetic Green's functions, and $G_\pm$ the corresponding scalar Green's functions.

To prove Eq. (16.21) we differentiate Eq. (16.6) covariantly with respect to x^β, use Eq. (13.3) to work on the right-hand side, and invoke Ricci's identity to permute the ordering of the covariant derivatives on the left-hand side. After simplification and involvement of the Ricci-flat condition (which, together with the Bianchi identities, implies that $R_{\alpha\gamma\beta\delta}{}^{;\beta} = 0$), we arrive at the equation

$$\Box\big(-G^{\alpha\beta}_{\ \ \gamma'\delta';\beta}\big) = -4\pi g^{\alpha}_{\ (\gamma'}\partial_{\delta')}\delta_4(x,x'). \tag{16.23}$$

Because this is also the differential equation satisfied by $G^{\alpha}_{\ (\beta';\gamma')}$, and because the solutions are chosen to satisfy the same boundary conditions, we have established the validity of Eq. (16.21).

The identity of Eq. (16.22) follows simply from the fact that $g^{\gamma'\delta'}G^{\alpha\beta}_{\ \ \gamma'\delta'}$ and $g^{\alpha\beta}G$ satisfy the same tensorial wave equation in a Ricci-flat spacetime.

16.5 Singular and regular Green's functions

We shall now construct singular and regular Green's functions for the linearized gravitational field. The treatment here parallels closely what was presented in Secs. 14.5 and 15.5.

We begin by introducing the bitensor $H^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ with properties

H1: $H^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ satisfies the homogeneous wave equation,

$$\Box H^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x') + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}(x)H^{\gamma\delta}_{\ \ \gamma'\delta'}(x,x') = 0; \tag{16.24}$$

H2: $H^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ is symmetric in its indices and arguments,

$$H_{\gamma'\delta'\alpha\beta}(x',x) = H_{\alpha\beta\gamma'\delta'}(x,x'); \tag{16.25}$$

H3: $H^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ agrees with the retarded Green's function if x is in the chronological future of x',

$$H^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x') = G_{+\ \gamma'\delta'}^{\ \alpha\beta}(x,x') \qquad \text{when } x\in I^+(x'); \tag{16.26}$$

H4: $H^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x')$ agrees with the advanced Green's function if x is in the chronological past of x',

$$H^{\alpha\beta}_{\ \ \gamma'\delta'}(x,x') = G_{-\ \gamma'\delta'}^{\ \alpha\beta}(x,x') \qquad \text{when } x\in I^-(x'). \tag{16.27}$$

It is easy to prove that property H4 follows from H2, H3, and the reciprocity relation (16.18) satisfied by the retarded and advanced Green's functions. That such a bitensor exists can be argued along the same lines as those presented in Sec. 14.5.

Equipped with $H^{\alpha\beta}{}_{\gamma'\delta'}(x,x')$ we define the singular Green's function to be

$$G_{\mathrm{S}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') = \frac{1}{2}\Big[G_{+}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') + G_{-}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') - H^{\alpha\beta}{}_{\gamma'\delta'}(x,x')\Big]. \tag{16.28}$$

This comes with the properties

S1: $G_{\mathrm{S}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x')$ satisfies the inhomogeneous wave equation,

$$\Box G_{\mathrm{S}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') + 2R_{\gamma}{}^{\alpha}{}_{\delta}{}^{\beta}(x) G_{\mathrm{S}}{}^{\gamma\delta}{}_{\gamma'\delta'}(x,x') = -4\pi g^{(\alpha}{}_{\gamma'}(x,x') g^{\beta)}{}_{\delta'}(x,x')\delta_4(x,x'); \tag{16.29}$$

S2: $G_{\mathrm{S}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x')$ is symmetric in its indices and arguments,

$$G^{\mathrm{S}}_{\gamma'\delta'\alpha\beta}(x',x) = G^{\mathrm{S}}_{\alpha\beta\gamma'\delta'}(x,x'); \tag{16.30}$$

S3: $G_{\mathrm{S}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x')$ vanishes if x is in the chronological past or future of x',

$$G_{\mathrm{S}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') = 0 \qquad \text{when } x \in I^{\pm}(x'). \tag{16.31}$$

These can be established as consequences of H1–H4 and the properties of the retarded and advanced Green's functions.

The regular two-point function is then defined by

$$G_{\mathrm{R}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') = G_{+}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') - G_{\mathrm{S}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x'), \tag{16.32}$$

and it comes with the properties

R1: $G_{\mathrm{R}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x')$ satisfies the homogeneous wave equation,

$$\Box G_{\mathrm{R}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') + 2R_{\gamma}{}^{\alpha}{}_{\delta}{}^{\beta}(x) G_{\mathrm{R}}{}^{\gamma\delta}{}_{\gamma'\delta'}(x,x') = 0; \tag{16.33}$$

R2: $G_{\mathrm{R}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x')$ agrees with the retarded Green's function if x is in the chronological future of x',

$$G_{\mathrm{R}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') = G_{+}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') \qquad \text{when } x \in I^{+}(x'); \tag{16.34}$$

R3: $G_{\mathrm{R}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x')$ vanishes if x is in the chronological past of x',

$$G_{\mathrm{R}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') = 0 \qquad \text{when } x \in I^{-}(x'). \tag{16.35}$$

Those follow immediately from S1–S3 and the properties of the retarded Green's function.

When x is restricted to the normal convex neighbourhood of x', we have the explicit relations

$$H^{\alpha\beta}{}_{\gamma'\delta'}(x,x') = V^{\alpha\beta}{}_{\gamma'\delta'}(x,x'), \tag{16.36}$$

$$G_{\mathrm{S}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') = \frac{1}{2}U^{\alpha\beta}{}_{\gamma'\delta'}(x,x')\delta(\sigma) - \frac{1}{2}V^{\alpha\beta}{}_{\gamma'\delta'}(x,x')\theta(\sigma), \tag{16.37}$$

$$G_{\mathrm{R}}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x') = \frac{1}{2}U^{\alpha\beta}{}_{\gamma'\delta'}(x,x')\Big[\delta_{+}(\sigma) - \delta_{-}(\sigma)\Big] + V^{\alpha\beta}{}_{\gamma'\delta'}(x,x')\Big[\theta_{+}(-\sigma) + \frac{1}{2}\theta(\sigma)\Big]. \tag{16.38}$$

From these we see clearly that the singular Green's function does not distinguish between past and future (property S2), and that its support excludes $I^{\pm}(x')$ (property S3). We see also that the regular two-point function coincides with $G_{+}{}^{\alpha\beta}{}_{\gamma'\delta'}(x,x')$ in $I^{+}(x')$ (property R2), and that its support does not include $I^{-}(x')$ (property R3).

Part IV
Motion of point particles

17 Motion of a scalar charge

17.1 Dynamics of a point scalar charge

A point particle carries a scalar charge q and moves on a world line γ described by relations $z^\mu(\lambda)$, in which λ is an arbitrary parameter. The particle generates a scalar potential $\Phi(x)$ and a field $\Phi_\alpha(x) := \nabla_\alpha \Phi(x)$. The dynamics of the entire system is governed by the action

$$S = S_{\text{field}} + S_{\text{particle}} + S_{\text{interaction}}, \tag{17.1}$$

where S_{field} is an action functional for a free scalar field in a spacetime with metric $g_{\alpha\beta}$, S_{particle} is the action of a free particle moving on a world line γ in this spacetime, and $S_{\text{interaction}}$ is an interaction term that couples the field to the particle.

The field action is given by

$$S_{\text{field}} = -\frac{1}{8\pi}\int \bigl(g^{\alpha\beta}\Phi_\alpha \Phi_\beta + \xi R \Phi^2\bigr)\sqrt{-g}\, d^4x, \tag{17.2}$$

where the integration is over all of spacetime; the field is coupled to the Ricci scalar R by an arbitrary constant ξ. The particle action is

$$S_{\text{particle}} = -m_0 \int_\gamma d\tau, \tag{17.3}$$

where m_0 is the bare mass of the particle and $d\tau = \sqrt{-g_{\mu\nu}(z)\dot{z}^\mu \dot{z}^\nu}\, d\lambda$ is the differential of proper time along the world line; we use an overdot on $z^\mu(\lambda)$ to indicate differentiation with respect to the parameter λ. Finally, the interaction term is given by

$$S_{\text{interaction}} = q\int_\gamma \Phi(z)\, d\tau = q\int \Phi(x)\delta_4(x,z)\sqrt{-g}\, d^4x d\tau. \tag{17.4}$$

Notice that both S_{particle} and $S_{\text{interaction}}$ are invariant under a reparameterization $\lambda \to \lambda'(\lambda)$ of the world line.

Demanding that the total action be stationary under a variation $\delta\Phi(x)$ of the field configuration yields the wave equation

$$\bigl(\Box - \xi R\bigr)\Phi(x) = -4\pi\mu(x) \tag{17.5}$$

for the scalar potential, with a charge density $\mu(x)$ defined by

$$\mu(x) = q\int_\gamma \delta_4(x,z)\, d\tau. \tag{17.6}$$

These equations determine the field $\Phi_\alpha(x)$ once the motion of the scalar charge is specified. On the other hand, demanding that the total action be stationary under a variation $\delta z^\mu(\lambda)$ of the world line yields the equations of motion

$$m(\tau)\frac{Du^\mu}{d\tau} = q\bigl(g^{\mu\nu} + u^\mu u^\nu\bigr)\Phi_\nu(z) \tag{17.7}$$

for the scalar charge. We have here adopted τ as the parameter on the world line, and introduced the four-velocity $u^\mu(\tau) := dz^\mu/d\tau$. The dynamical mass that appears in Eq. (17.7) is defined by $m(\tau) := m_0 - q\Phi(z)$, which can also be expressed in differential form as

$$\frac{dm}{d\tau} = -q\Phi_\mu(z)u^\mu. \tag{17.8}$$

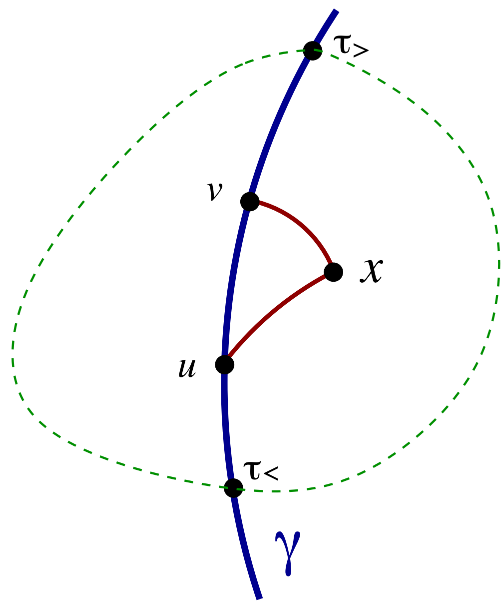

Figure 9: The region within the dashed boundary represents the normal convex neighbourhood of the point x. The world line γ enters the neighbourhood at proper time $\tau_<$ and exits at proper time $\tau_>$. Also shown are the retarded point $z(u)$ and the advanced point $z(v)$.

It should be clear that Eqs. (17.7) and (17.8) are valid only in a formal sense, because the scalar potential obtained from Eqs. (17.5) and (17.6) diverges on the world line. Before we can make sense of these equations we have to analyze the field's singularity structure near the world line.

17.2 Retarded potential near the world line

The retarded solution to Eq. (17.5) is $\Phi(x) = \int G_+(x,x')\mu(x')\sqrt{g'}\,d^4x'$, where $G_+(x,x')$ is the retarded Green's function introduced in Sec. 14. After substitution of Eq. (17.6) we obtain

$$\Phi(x) = q\int_\gamma G_+(x,z)\,d\tau, \tag{17.9}$$

in which $z(\tau)$ gives the description of the world line γ. Because the retarded Green's function is defined globally in the entire spacetime, Eq. (17.9) applies to any field point x.

We now specialize Eq. (17.9) to a point x near the world line; see Fig. 9. We let $\mathcal{N}(x)$ be the normal convex neighbourhood of this point, and we assume that the world line traverses $\mathcal{N}(x)$. Let $\tau_<$ be the value of the proper-time parameter at which γ enters $\mathcal{N}(x)$ from the past, and let $\tau_>$ be its value when the world line leaves $\mathcal{N}(x)$. Then Eq. (17.9) can be broken up into the three integrals

$$\Phi(x) = q\int_{-\infty}^{\tau_<} G_+(x,z)\,d\tau + q\int_{\tau_<}^{\tau_>} G_+(x,z)\,d\tau + q\int_{\tau_>}^{\infty} G_+(x,z)\,d\tau.$$

The third integration vanishes because x is then in the past of $z(\tau)$, and $G_+(x,z) = 0$. For the second integration, x is the normal convex neighbourhood of $z(\tau)$, and the retarded Green's function can be expressed in the Hadamard form produced in Sec. 14.2. This gives

$$\int_{\tau_<}^{\tau_>} G_+(x,z)\,d\tau = \int_{\tau_<}^{\tau_>} U(x,z)\delta_+(\sigma)\,d\tau + \int_{\tau_<}^{\tau_>} V(x,z)\theta_+(-\sigma)\,d\tau,$$

and to evaluate this we refer back to Sec. 10 and let $x' := z(u)$ be the retarded point associated with x; these points are related by $\sigma(x,x') = 0$ and $r := \sigma_{\alpha'}u^{\alpha'}$ is the retarded distance between x and the world line. We resume the index convention of Sec. 10: to tensors at x we assign indices α, β, etc.; to tensors at

x' we assign indices α', β', etc.; and to tensors at a generic point $z(\tau)$ on the world line we assign indices μ, ν, etc.

To perform the first integration we change variables from τ to σ, noticing that σ increases as $z(\tau)$ passes through x'. The change of σ on the world line is given by $d\sigma := \sigma(x, z + dz) - \sigma(x, z) = \sigma_\mu u^\mu \, d\tau$, and we find that the first integral evaluates to $U(x, z)/(\sigma_\mu u^\mu)$ with z identified with x'. The second integration is cut off at $\tau = u$ by the step function, and we obtain our final expression for the retarded potential of a point scalar charge:

$$\Phi(x) = \frac{q}{r} U(x, x') + q \int_{\tau_<}^{u} V(x, z)\, d\tau + q \int_{-\infty}^{\tau_<} G_+(x, z)\, d\tau. \tag{17.10}$$

This expression applies to a point x sufficiently close to the world line that there exists a nonempty intersection between $\mathcal{N}(x)$ and γ.

17.3 Field of a scalar charge in retarded coordinates

When we differentiate the potential of Eq. (17.10) we must keep in mind that a variation in x induces a variation in x' because the new points $x + \delta x$ and $x' + \delta x'$ must also be linked by a null geodesic — you may refer back to Sec. 10.2 for a detailed discussion. This means, for example, that the total variation of $U(x, x')$ is $\delta U = U(x + \delta x, x' + \delta x') - U(x, x') = U_{;\alpha} \delta x^\alpha + U_{;\alpha'} u^{\alpha'} \, \delta u$. The gradient of the scalar potential is therefore given by

$$\Phi_\alpha(x) = -\frac{q}{r^2} U(x, x') \partial_\alpha r + \frac{q}{r} U_{;\alpha}(x, x') + \frac{q}{r} U_{;\alpha'}(x, x') u^{\alpha'} \partial_\alpha u + q V(x, x') \partial_\alpha u + \Phi^{\rm tail}_\alpha(x), \tag{17.11}$$

where the "tail integral" is defined by

$$\begin{aligned} \Phi^{\rm tail}_\alpha(x) &= q \int_{\tau_<}^{u} \nabla_\alpha V(x, z)\, d\tau + q \int_{-\infty}^{\tau_<} \nabla_\alpha G_+(x, z)\, d\tau \\ &= q \int_{-\infty}^{u^-} \nabla_\alpha G_+(x, z)\, d\tau. \end{aligned} \tag{17.12}$$

In the second form of the definition we integrate $\nabla_\alpha G_+(x, z)$ from $\tau = -\infty$ to almost $\tau = u$, but we cut the integration short at $\tau = u^- := u - 0^+$ to avoid the singular behaviour of the retarded Green's function at $\sigma = 0$. This limiting procedure gives rise to the first form of the definition, with the advantage that the integral need not be broken up into contributions that refer to $\mathcal{N}(x)$ and its complement, respectively.

We shall now expand $\Phi_\alpha(x)$ in powers of r, and express the results in terms of the retarded coordinates (u, r, Ω^a) introduced in Sec. 10. It will be convenient to decompose $\Phi_\alpha(x)$ in the tetrad (e^α_0, e^α_a) that is obtained by parallel transport of $(u^{\alpha'}, e^{\alpha'}_a)$ on the null geodesic that links x to $x' := z(u)$; this construction is detailed in Sec. 10. The expansion relies on Eq. (10.29) for $\partial_\alpha u$, Eq. (10.31) for $\partial_\alpha r$, and we shall need

$$U(x, x') = 1 + \frac{1}{12} r^2 \bigl(R_{00} + 2R_{0a}\Omega^a + R_{ab}\Omega^a\Omega^b \bigr) + O(r^3), \tag{17.13}$$

which follows from Eq. (14.10) and the relation $\sigma^{\alpha'} = -r(u^{\alpha'} + \Omega^a e^{\alpha'}_a)$ first encountered in Eq. (10.7); recall that

$$R_{00}(u) = R_{\alpha'\beta'} u^{\alpha'} u^{\beta'}, \qquad R_{0a}(u) = R_{\alpha'\beta'} u^{\alpha'} e^{\beta'}_a, \qquad R_{ab}(u) = R_{\alpha'\beta'} e^{\alpha'}_a e^{\beta'}_b$$

are frame components of the Ricci tensor evaluated at x'. We shall also need the expansions

$$U_{;\alpha}(x, x') = \frac{1}{6} r g^{\alpha'}_{\ \alpha} \bigl(R_{\alpha'0} + R_{\alpha'b}\Omega^b \bigr) + O(r^2) \tag{17.14}$$

and

$$U_{;\alpha'}(x, x') u^{\alpha'} = -\frac{1}{6} r \bigl(R_{00} + R_{0a}\Omega^a \bigr) + O(r^2) \tag{17.15}$$

which follow from Eqs. (14.11); recall from Eq. (10.4) that the parallel propagator can be expressed as $g^{\alpha'}_{\ \alpha} = u^{\alpha'} e^0_\alpha + e^{\alpha'}_a e^a_\alpha$. And finally, we shall need

$$V(x, x') = \frac{1}{12} \bigl(1 - 6\xi \bigr) R + O(r), \tag{17.16}$$

a relation that was first established in Eq. (14.13); here $R := R(u)$ is the Ricci scalar evaluated at x'.

Collecting all these results gives

$$\begin{aligned}
\Phi_0(u,r,\Omega^a) &:= \Phi_\alpha(x) e^\alpha_0(x) \\
&= \frac{q}{r} a_a \Omega^a + \frac{1}{2} q R_{a0b0}\Omega^a\Omega^b + \frac{1}{12}\bigl(1-6\xi\bigr) qR + \Phi^{\rm tail}_0 + O(r), \qquad (17.17)\\
\Phi_a(u,r,\Omega^a) &:= \Phi_\alpha(x) e^\alpha_a(x) \\
&= -\frac{q}{r^2}\Omega_a - \frac{q}{r} a_b\Omega^b\Omega_a - \frac{1}{3} q R_{b0c0}\Omega^b\Omega^c\Omega_a - \frac{1}{6} q\bigl(R_{a0b0}\Omega^b - R_{ab0c}\Omega^b\Omega^c\bigr) \\
&\quad + \frac{1}{12} q\bigl[R_{00} - R_{bc}\Omega^b\Omega^c - (1-6\xi)R\bigr]\Omega_a + \frac{1}{6} q\bigl(R_{a0} + R_{ab}\Omega^b\bigr) + \Phi^{\rm tail}_a + O(r), \qquad (17.18)
\end{aligned}$$

where $a_a = a_{\alpha'} e^{\alpha'}_a$ are the frame components of the acceleration vector,

$$R_{a0b0}(u) = R_{\alpha'\gamma'\beta'\delta'} e^{\alpha'}_a u^{\gamma'} e^{\beta'}_b u^{\delta'}, \qquad R_{ab0c}(u) = R_{\alpha'\gamma'\beta'\delta'} e^{\alpha'}_a e^{\gamma'}_b u^{\beta'} e^{\delta'}_c$$

are frame components of the Riemann tensor evaluated at x', and

$$\Phi^{\rm tail}_0(u) = \Phi^{\rm tail}_{\alpha'}(x') u^{\alpha'}, \qquad \Phi^{\rm tail}_a(u) = \Phi^{\rm tail}_{\alpha'}(x') e^{\alpha'}_a \qquad (17.19)$$

are the frame components of the tail integral evaluated at x'. Equations (17.17) and (17.18) show clearly that $\Phi_\alpha(x)$ is singular on the world line: the field diverges as r^{-2} when $r \to 0$, and many of the terms that stay bounded in the limit depend on Ω^a and therefore possess a directional ambiguity at $r = 0$.

17.4 Field of a scalar charge in Fermi normal coordinates

The gradient of the scalar potential can also be expressed in the Fermi normal coordinates of Sec. 9. To effect this translation we make $\bar{x} := z(t)$ the new reference point on the world line. We resume here the notation of Sec. 11 and assign indices $\bar\alpha$, $\bar\beta$, ... to tensors at $\bar{x}$. The Fermi normal coordinates are denoted (t, s, ω^a), and we let $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ be the tetrad at x that is obtained by parallel transport of $(u^{\bar\alpha}, e^{\bar\alpha}_a)$ on the spacelike geodesic that links x to $\bar{x}$.

Our first task is to decompose $\Phi_\alpha(x)$ in the tetrad $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$, thereby defining $\bar\Phi_0 := \Phi_\alpha \bar{e}^\alpha_0$ and $\bar\Phi_a := \Phi_\alpha \bar{e}^\alpha_a$. For this purpose we use Eqs. (11.7), (11.8), (17.17), and (17.18) to obtain

$$\begin{aligned}
\bar\Phi_0 &= \Bigl[1 + O(r^2)\Bigr]\Phi_0 + \Bigl[r\bigl(1 - a_b\Omega^b\bigr)a^a + \frac{1}{2} r^2 \dot{a}^a + \frac{1}{2} r^2 R^a_{\ 0b0}\Omega^b + O(r^3)\Bigr]\Phi_a \\
&= -\frac{1}{2} q \dot{a}_a \Omega^a + \frac{1}{12}(1-6\xi) qR + \bar\Phi^{\rm tail}_0 + O(r)
\end{aligned}$$

and

$$\begin{aligned}
\bar\Phi_a &= \Bigl[\delta^b_{\ a} + \frac{1}{2} r^2 a^b a_a - \frac{1}{2} r^2 R^b_{\ a0c}\Omega^c + O(r^3)\Bigr]\Phi_b + \Bigl[r a_a + O(r^2)\Bigr]\Phi_0 \\
&= -\frac{q}{r^2}\Omega_a - \frac{q}{r} a_b\Omega^b\Omega_a + \frac{1}{2} q a_b\Omega^b a_a - \frac{1}{3} q R_{b0c0}\Omega^b\Omega^c\Omega_a - \frac{1}{6} q R_{a0b0}\Omega^b - \frac{1}{3} q R_{ab0c}\Omega^b\Omega^c \\
&\quad + \frac{1}{12} q\bigl[R_{00} - R_{bc}\Omega^b\Omega^c - (1-6\xi)R\bigr]\Omega_a + \frac{1}{6} q\bigl(R_{a0} + R_{ab}\Omega^b\bigr) + \bar\Phi^{\rm tail}_a + O(r),
\end{aligned}$$

where all frame components are still evaluated at x', except for $\bar\Phi^{\rm tail}_0$ and $\bar\Phi^{\rm tail}_a$ which are evaluated at $\bar{x}$.

We must still translate these results into the Fermi normal coordinates (t, s, ω^a). For this we involve Eqs. (11.4), (11.5), and (11.6), from which we deduce, for example,

$$\begin{aligned}
\frac{1}{r^2}\Omega_a &= \frac{1}{s^2}\omega_a + \frac{1}{2s} a_a - \frac{3}{2s} a_b\omega^b\omega_a - \frac{3}{4} a_b\omega^b a_a + \frac{15}{8}\bigl(a_b\omega^b\bigr)^2\omega_a + \frac{3}{8}\dot{a}_0\omega_a - \frac{1}{3}\dot{a}_a \\
&\quad + \dot{a}_b\omega^b\omega_a + \frac{1}{6} R_{a0b0}\omega^b - \frac{1}{2} R_{b0c0}\omega^b\omega^c\omega_a - \frac{1}{3} R_{ab0c}\omega^b\omega^c + O(s)
\end{aligned}$$

and

$$\frac{1}{r} a_b\Omega^b\Omega_a = \frac{1}{s} a_b\omega^b\omega_a + \frac{1}{2} a_b\omega^b a_a - \frac{3}{2}\bigl(a_b\omega^b\bigr)^2\omega_a - \frac{1}{2}\dot{a}_0\omega_a - \dot{a}_b\omega^b\omega_a + O(s),$$

in which all frame components (on the right-hand side of these relations) are now evaluated at $\bar{x}$; to obtain the second relation we expressed $a_a(u)$ as $a_a(t) - s\dot{a}_a(t) + O(s^2)$ since according to Eq. (11.4), $u = t - s + O(s^2)$.

Collecting these results yields

$$\begin{aligned} \bar{\Phi}_0(t, s, \omega^a) &:= \Phi_\alpha(x)\bar{e}_0^\alpha(x) \\ &= -\frac{1}{2}q\dot{a}_a\omega^a + \frac{1}{12}(1-6\xi)qR + \bar{\Phi}_0^{\rm tail} + O(s), \end{aligned} \tag{17.20}$$

$$\begin{aligned} \bar{\Phi}_a(t, s, \omega^a) &:= \Phi_\alpha(x)\bar{e}_a^\alpha(x) \\ &= -\frac{q}{s^2}\omega_a - \frac{q}{2s}\bigl(a_a - a_b\omega^b\omega_a\bigr) + \frac{3}{4}qa_b\omega^b a_a - \frac{3}{8}q\bigl(a_b\omega^b\bigr)^2\omega_a + \frac{1}{8}q\dot{a}_0\omega_a + \frac{1}{3}q\dot{a}_a \\ &\quad - \frac{1}{3}qR_{a0b0}\omega^b + \frac{1}{6}qR_{b0c0}\omega^b\omega^c\omega_a + \frac{1}{12}q\bigl[R_{00} - R_{bc}\omega^b\omega^c - (1-6\xi)R\bigr]\omega_a \\ &\quad + \frac{1}{6}q\bigl(R_{a0} + R_{ab}\omega^b\bigr) + \bar{\Phi}_a^{\rm tail} + O(s). \end{aligned} \tag{17.21}$$

In these expressions, $a_a(t) = a_{\bar{\alpha}}e_a^{\bar{\alpha}}$ are the frame components of the acceleration vector evaluated at $\bar{x}$, $\dot{a}_0(t) = \dot{a}_{\bar{\alpha}}u^{\bar{\alpha}}$ and $\dot{a}_a(t) = \dot{a}_{\bar{\alpha}}e_a^{\bar{\alpha}}$ are frame components of its covariant derivative, $R_{a0b0}(t) = R_{\bar{\alpha}\bar{\gamma}\bar{\beta}\bar{\delta}}e_a^{\bar{\alpha}}u^{\bar{\gamma}}e_b^{\bar{\beta}}u^{\bar{\delta}}$ are frame components of the Riemann tensor evaluated at $\bar{x}$,

$$R_{00}(t) = R_{\bar{\alpha}\bar{\beta}}u^{\bar{\alpha}}u^{\bar{\beta}}, \qquad R_{0a}(t) = R_{\bar{\alpha}\bar{\beta}}u^{\bar{\alpha}}e_a^{\bar{\beta}}, \qquad R_{ab}(t) = R_{\bar{\alpha}\bar{\beta}}e_a^{\bar{\alpha}}e_b^{\bar{\beta}}$$

are frame components of the Ricci tensor, and $R(t)$ is the Ricci scalar evaluated at $\bar{x}$. Finally, we have that

$$\bar{\Phi}_0^{\rm tail}(t) = \Phi_{\bar{\alpha}}^{\rm tail}(\bar{x})u^{\bar{\alpha}}, \qquad \bar{\Phi}_a^{\rm tail}(t) = \Phi_{\bar{\alpha}}^{\rm tail}(\bar{x})e_a^{\bar{\alpha}} \tag{17.22}$$

are the frame components of the tail integral — see Eq. (17.12) — evaluated at $\bar{x} := z(t)$.

We shall now compute the averages of $\bar{\Phi}_0$ and $\bar{\Phi}_a$ over $S(t,s)$, a two-surface of constant t and s; these will represent the mean value of the field at a fixed proper distance away from the world line, as measured in a reference frame that is momentarily comoving with the particle. The two-surface is charted by angles θ^A $(A = 1, 2)$ and it is described, in the Fermi normal coordinates, by the parametric relations $\hat{x}^a = s\omega^a(\theta^A)$; a canonical choice of parameterization is $\omega^a = (\sin\theta\cos\phi, \sin\theta\sin\phi, \cos\theta)$. Introducing the transformation matrices $\omega_A^a := \partial\omega^a/\partial\theta^A$, we find from Eq. (9.16) that the induced metric on $S(t,s)$ is given by

$$ds^2 = s^2\Bigl[\omega_{AB} - \frac{1}{3}s^2 R_{AB} + O(s^3)\Bigr]\,d\theta^A d\theta^B, \tag{17.23}$$

where $\omega_{AB} := \delta_{ab}\omega_A^a\omega_B^b$ is the metric of the unit two-sphere, and where $R_{AB} := R_{acbd}\omega_A^a\omega^c\omega_B^b\omega^d$ depends on t and the angles θ^A. From this we infer that the element of surface area is given by

$$d\mathcal{A} = s^2\Bigl[1 - \frac{1}{6}s^2 R^c{}_{acb}(t)\omega^a\omega^b + O(s^3)\Bigr]\,d\omega, \tag{17.24}$$

where $d\omega = \sqrt{\det[\omega_{AB}]}\,d^2\theta$ is an element of solid angle — in the canonical parameterization, $d\omega = \sin\theta\,d\theta d\phi$. Integration of Eq. (17.24) produces the total surface area of $S(t,s)$, and $\mathcal{A} = 4\pi s^2[1 - \frac{1}{18}s^2 R^{ab}{}_{ab} + O(s^3)]$.

The averaged fields are defined by

$$\bigl\langle\bar{\Phi}_0\bigr\rangle(t,s) = \frac{1}{\mathcal{A}}\oint_{S(t,s)}\bar{\Phi}_0(t,s,\theta^A)\,d\mathcal{A}, \qquad \bigl\langle\bar{\Phi}_a\bigr\rangle(t,s) = \frac{1}{\mathcal{A}}\oint_{S(t,s)}\bar{\Phi}_a(t,s,\theta^A)\,d\mathcal{A}, \tag{17.25}$$

where the quantities to be integrated are scalar functions of the Fermi normal coordinates. The results

$$\frac{1}{4\pi}\oint\omega^a\,d\omega = 0, \qquad \frac{1}{4\pi}\oint\omega^a\omega^b\,d\omega = \frac{1}{3}\delta^{ab}, \qquad \frac{1}{4\pi}\oint\omega^a\omega^b\omega^c\,d\omega = 0, \tag{17.26}$$

are easy to establish, and we obtain

$$\bigl\langle\bar{\Phi}_0\bigr\rangle = \frac{1}{12}(1-6\xi)qR + \bar{\Phi}_0^{\rm tail} + O(s), \tag{17.27}$$

$$\bigl\langle\bar{\Phi}_a\bigr\rangle = -\frac{q}{3s}a_a + \frac{1}{3}q\dot{a}_a + \frac{1}{6}qR_{a0} + \bar{\Phi}_a^{\rm tail} + O(s). \tag{17.28}$$

The averaged field is still singular on the world line. Regardless, we shall take the formal limit $s \to 0$ of the expressions displayed in Eqs. (17.27) and (17.28). In the limit the tetrad $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ reduces to $(u^{\bar{\alpha}}, e^{\bar{\alpha}}_a)$, and we can reconstruct the field at $\bar{x}$ by invoking the completeness relations $\delta^{\bar{\alpha}}_{\bar{\beta}} = -u^{\bar{\alpha}} u_{\bar{\beta}} + e^{\bar{\alpha}}_a e^a_{\bar{\beta}}$. We thus obtain

$$\big\langle \Phi_{\bar{\alpha}} \big\rangle = \lim_{s\to 0}\biggl(-\frac{q}{3s}\biggr) a_{\bar{\alpha}} - \frac{1}{12}(1-6\xi) q R u_{\bar{\alpha}} + q\bigl(g_{\bar{\alpha}\bar{\beta}} + u_{\bar{\alpha}} u_{\bar{\beta}}\bigr)\biggl(\frac{1}{3}\dot{a}^{\bar{\beta}} + \frac{1}{6} R^{\bar{\beta}}{}_{\bar{\gamma}} u^{\bar{\gamma}}\biggr) + \Phi^{\rm tail}_{\bar{\alpha}}, \tag{17.29}$$

where the tail integral can be copied from Eq. (17.12),

$$\Phi^{\rm tail}_{\bar{\alpha}}(\bar{x}) = q \int_{-\infty}^{t^-} \nabla_{\bar{\alpha}} G_+(\bar{x}, z)\, d\tau. \tag{17.30}$$

The tensors appearing in Eq. (17.29) all refer to $\bar{x} := z(t)$, which now stands for an arbitrary point on the world line γ.

17.5 Singular and regular fields

The singular potential

$$\Phi^{\rm S}(x) = q \int_\gamma G_{\rm S}(x,z)\, d\tau \tag{17.31}$$

is the (unphysical) solution to Eqs. (17.5) and (17.6) that is obtained by adopting the singular Green's function of Eq. (14.30) instead of the retarded Green's function. As we shall see, the resulting singular field $\Phi^{\rm S}_\alpha(x)$ reproduces the singular behaviour of the retarded solution; the difference, $\Phi^{\rm R}_\alpha(x) = \Phi_\alpha(x) - \Phi^{\rm S}_\alpha(x)$, is regular on the world line.

To evaluate the integral of Eq. (17.31) we assume once more that x is sufficiently close to γ that the world line traverses $\mathcal{N}(x)$; refer back to Fig. 9. As before we let $\tau_<$ and $\tau_>$ be the values of the proper-time parameter at which γ enters and leaves $\mathcal{N}(x)$, respectively. Then Eq. (17.31) can be broken up into the three integrals

$$\Phi^{\rm S}(x) = q \int_{-\infty}^{\tau_<} G_{\rm S}(x,z)\, d\tau + q \int_{\tau_<}^{\tau_>} G_{\rm S}(x,z)\, d\tau + q \int_{\tau_>}^{\infty} G_{\rm S}(x,z)\, d\tau.$$

The first integration vanishes because x is then in the chronological future of $z(\tau)$, and $G_{\rm S}(x,z) = 0$ by Eq. (14.21). Similarly, the third integration vanishes because x is then in the chronological past of $z(\tau)$. For the second integration, x is the normal convex neighbourhood of $z(\tau)$, the singular Green's function can be expressed in the Hadamard form of Eq. (14.32), and we have

$$\int_{\tau_<}^{\tau_>} G_{\rm S}(x,z)\, d\tau = \frac{1}{2}\int_{\tau_<}^{\tau_>} U(x,z)\delta_+(\sigma)\, d\tau + \frac{1}{2}\int_{\tau_<}^{\tau_>} U(x,z)\delta_-(\sigma)\, d\tau - \frac{1}{2}\int_{\tau_<}^{\tau_>} V(x,z)\theta(\sigma)\, d\tau.$$

To evaluate these we re-introduce the retarded point $x' := z(u)$ and let $x'' := z(v)$ be the *advanced point* associated with x; we recall from Sec. 11.4 that these points are related by $\sigma(x,x'') = 0$ and that $r_{\rm adv} := -\sigma_{\alpha''} u^{\alpha''}$ is the advanced distance between x and the world line.

To perform the first integration we change variables from τ to σ, noticing that σ increases as $z(\tau)$ passes through x'; the integral evaluates to $U(x,x')/r$. We do the same for the second integration, but we notice now that σ decreases as $z(\tau)$ passes through x''; the integral evaluates to $U(x,x'')/r_{\rm adv}$. The third integration is restricted to the interval $u \le \tau \le v$ by the step function, and we obtain our final expression for the singular potential of a point scalar charge:

$$\Phi^{\rm S}(x) = \frac{q}{2r} U(x,x') + \frac{q}{2r_{\rm adv}} U(x,x'') - \frac{1}{2} q \int_u^v V(x,z)\, d\tau. \tag{17.32}$$

We observe that $\Phi^{\rm S}(x)$ depends on the state of motion of the scalar charge between the retarded time u and the advanced time v; contrary to what was found in Sec. 17.2 for the retarded potential, there is no dependence on the particle's remote past.

We use the techniques of Sec. 17.3 to differentiate the potential of Eq. (17.32). We find

$$\begin{aligned}\Phi^{\rm S}_\alpha(x) &= -\frac{q}{2r^2}U(x,x')\partial_\alpha r - \frac{q}{2{r_{\rm adv}}^2}U(x,x'')\partial_\alpha r_{\rm adv} + \frac{q}{2r}U_{;\alpha}(x,x') + \frac{q}{2r}U_{;\alpha'}(x,x')u^{\alpha'}\partial_\alpha u \\ &\quad + \frac{q}{2r_{\rm adv}}U_{;\alpha}(x,x'') + \frac{q}{2r_{\rm adv}}U_{;\alpha''}(x,x'')u^{\alpha''}\partial_\alpha v + \frac{1}{2}qV(x,x')\partial_\alpha u - \frac{1}{2}qV(x,x'')\partial_\alpha v \\ &\quad - \frac{1}{2}q\int_u^v \nabla_\alpha V(x,z)\,d\tau, \end{aligned} \tag{17.33}$$

and we would like to express this as an expansion in powers of r. For this we shall rely on results already established in Sec. 17.3, as well as additional expansions that will involve the advanced point x''. Those we develop now.

We recall first that a relation between retarded and advanced times was worked out in Eq. (11.12), that an expression for the advanced distance was displayed in Eq. (11.13), and that Eqs. (11.14) and (11.15) give expansions for $\partial_\alpha v$ and $\partial_\alpha r_{\rm adv}$, respectively.

To derive an expansion for $U(x,x'')$ we follow the general method of Sec. 11.4 and define a function $U(\tau) := U(x,z(\tau))$ of the proper-time parameter on γ. We have that

$$U(x,x'') := U(v) = U(u+\Delta') = U(u) + \dot{U}(u)\Delta' + \frac{1}{2}\ddot{U}(u)\Delta'^2 + O\big(\Delta'^3\big),$$

where overdots indicate differentiation with respect to τ, and where $\Delta' := v-u$. The leading term $U(u) := U(x,x')$ was worked out in Eq. (17.13), and the derivatives of $U(\tau)$ are given by

$$\dot{U}(u) = U_{;\alpha'}u^{\alpha'} = -\frac{1}{6}r\big(R_{00} + R_{0a}\Omega^a\big) + O(r^2)$$

and

$$\ddot{U}(u) = U_{;\alpha'\beta'}u^{\alpha'}u^{\beta'} + U_{;\alpha'}a^{\alpha'} = \frac{1}{6}R_{00} + O(r),$$

according to Eqs. (17.15) and (14.11). Combining these results together with Eq. (11.12) for Δ' gives

$$U(x,x'') = 1 + \frac{1}{12}r^2\big(R_{00} - 2R_{0a}\Omega^a + R_{ab}\Omega^a\Omega^b\big) + O(r^3), \tag{17.34}$$

which should be compared with Eq. (17.13). It should be emphasized that in Eq. (17.34) and all equations below, the frame components of the Ricci tensor are evaluated at the retarded point $x' := z(u)$, and not at the advanced point. The preceding computation gives us also an expansion for $U_{;\alpha''}u^{\alpha''} := \dot{U}(v) = \dot{U}(u) + \ddot{U}(u)\Delta' + O(\Delta'^2)$. This becomes

$$U_{;\alpha''}(x,x'')u^{\alpha''} = \frac{1}{6}r\big(R_{00} - R_{0a}\Omega^a\big) + O(r^2), \tag{17.35}$$

which should be compared with Eq. (17.15).

We proceed similarly to derive an expansion for $U_{;\alpha}(x,x'')$. Here we introduce the functions $U_\alpha(\tau) := U_{;\alpha}(x,z(\tau))$ and express $U_{;\alpha}(x,x'')$ as $U_\alpha(v) = U_\alpha(u) + \dot{U}_\alpha(u)\Delta' + O(\Delta'^2)$. The leading term $U_\alpha(u) := U_{;\alpha}(x,x')$ was computed in Eq. (17.14), and

$$\dot{U}_\alpha(u) = U_{;\alpha\beta'}u^{\beta'} = -\frac{1}{6}g^{\alpha'}_{\ \alpha}R_{\alpha'0} + O(r)$$

follows from Eq. (14.11). Combining these results together with Eq. (11.12) for Δ' gives

$$U_{;\alpha}(x,x'') = -\frac{1}{6}rg^{\alpha'}_{\ \alpha}\Big(R_{\alpha'0} - R_{\alpha'b}\Omega^b\Big) + O(r^2), \tag{17.36}$$

and this should be compared with Eq. (17.14).

The last expansion we shall need is

$$V(x,x'') = \frac{1}{12}\big(1-6\xi\big)R + O(r), \tag{17.37}$$

which follows at once from Eq. (17.16) and the fact that $V(x,x'') - V(x,x') = O(r)$; the Ricci scalar is evaluated at the retarded point x'.

It is now a straightforward (but tedious) matter to substitute these expansions (all of them!) into Eq. (17.33) and obtain the projections of the singular field $\Phi^{\rm S}_\alpha(x)$ in the same tetrad (e^α_0, e^α_a) that was employed in Sec. 17.3. This gives

$$\begin{aligned}
\Phi^{\rm S}_0(u,r,\Omega^a) &:= \Phi^{\rm S}_\alpha(x) e^\alpha_0(x) \\
&= \frac{q}{r} a_a \Omega^a + \frac{1}{2} q R_{a0b0} \Omega^a \Omega^b + O(r), \qquad (17.38)\\
\Phi^{\rm S}_a(u,r,\Omega^a) &:= \Phi^{\rm S}_\alpha(x) e^\alpha_a(x) \\
&= -\frac{q}{r^2}\Omega_a - \frac{q}{r} a_b \Omega^b \Omega_a - \frac{1}{3} q \dot{a}_a - \frac{1}{3} q R_{b0c0}\Omega^b\Omega^c\Omega_a - \frac{1}{6} q \bigl(R_{a0b0}\Omega^b - R_{ab0c}\Omega^b\Omega^c \bigr) \\
&\quad + \frac{1}{12} q \bigl[R_{00} - R_{bc}\Omega^b\Omega^c - (1-6\xi)R \bigr] \Omega_a + \frac{1}{6} q R_{ab}\Omega^b, \qquad (17.39)
\end{aligned}$$

in which all frame components are evaluated at the retarded point $x' := z(u)$. Comparison of these expressions with Eqs. (17.17) and (17.18) reveals that the retarded and singular fields share the same singularity structure.

The difference between the retarded field of Eqs. (17.17), (17.18) and the singular field of Eqs. (17.38), (17.39) defines the regular field $\Phi^{\rm R}_\alpha(x)$. Its frame components are

$$\Phi^{\rm R}_0 = \frac{1}{12}(1-6\xi) q R + \Phi^{\rm tail}_0 + O(r), \qquad (17.40)$$

$$\Phi^{\rm R}_a = \frac{1}{3} q \dot{a}_a + \frac{1}{6} q R_{a0} + \Phi^{\rm tail}_a + O(r), \qquad (17.41)$$

and we see that $\Phi^{\rm R}_\alpha(x)$ is a regular vector field on the world line. There is therefore no obstacle in evaluating the regular field directly at $x = x'$, where the tetrad (e^α_0, e^α_a) becomes $(u^{\alpha'}, e^{\alpha'}_a)$. Reconstructing the field at x' from its frame components, we obtain

$$\Phi^{\rm R}_{\alpha'}(x') = -\frac{1}{12}(1-6\xi) q R u_{\alpha'} + q \bigl(g_{\alpha'\beta'} + u_{\alpha'} u_{\beta'} \bigr) \biggl(\frac{1}{3} \dot{a}^{\beta'} + \frac{1}{6} R^{\beta'}_{\ \ \gamma'} u^{\gamma'} \biggr) + \Phi^{\rm tail}_{\alpha'}, \qquad (17.42)$$

where the tail term can be copied from Eq. (17.12),

$$\Phi^{\rm tail}_{\alpha'}(x') = q \int_{-\infty}^{u^-} \nabla_{\alpha'} G_+(x',z)\, d\tau. \qquad (17.43)$$

The tensors appearing in Eq. (17.42) all refer to the retarded point $x' := z(u)$, which now stands for an arbitrary point on the world line γ.

17.6 Equations of motion

The retarded field $\Phi_\alpha(x)$ of a point scalar charge is singular on the world line, and this behaviour makes it difficult to understand how the field is supposed to act on the particle and affect its motion. The field's singularity structure was analyzed in Secs. 17.3 and 17.4, and in Sec. 17.5 it was shown to originate from the singular field $\Phi^{\rm S}_\alpha(x)$; the regular field $\Phi^{\rm R}_\alpha(x) = \Phi_\alpha(x) - \Phi^{\rm S}_\alpha(x)$ was then shown to be regular on the world line.

To make sense of the retarded field's action on the particle we temporarily model the scalar charge not as a point particle, but as a small hollow shell that appears spherical when observed in a reference frame that is momentarily comoving with the particle; the shell's radius is s_0 in Fermi normal coordinates, and it is independent of the angles contained in the unit vector ω^a. The *net force* acting at proper time τ on this hollow shell is the average of $q\Phi_\alpha(\tau, s_0, \omega^a)$ over the surface of the shell. Assuming that the field on the shell is equal to the field of a point particle evaluated at $s = s_0$, and ignoring terms that disappear in the limit $s_0 \to 0$, we obtain from Eq. (17.29)

$$q \langle \Phi_\mu \rangle = -(\delta m) a_\mu - \frac{1}{12}(1-6\xi) q^2 R u_\mu + q^2 \bigl(g_{\mu\nu} + u_\mu u_\nu \bigr) \biggl(\frac{1}{3} \dot{a}^\nu + \frac{1}{6} R^\nu_{\ \lambda} u^\lambda \biggr) + q \Phi^{\rm tail}_\mu, \qquad (17.44)$$

where

$$\delta m := \lim_{s_0 \to 0} \frac{q^2}{3s_0} \tag{17.45}$$

is formally a divergent quantity and

$$q\Phi^{\rm tail}_\mu = q^2 \int_{-\infty}^{\tau^-} \nabla_\mu G_+\bigl(z(\tau), z(\tau')\bigr)\, d\tau' \tag{17.46}$$

is the tail part of the force; all tensors in Eq. (17.44) are evaluated at an arbitrary point $z(\tau)$ on the world line.

Substituting Eqs. (17.44) and (17.46) into Eq. (17.7) gives rise to the equations of motion

$$\bigl(m + \delta m\bigr) a^\mu = q^2 \bigl(\delta^\mu_{\ \nu} + u^\mu u_\nu\bigr) \biggl[\frac{1}{3}\dot{a}^\nu + \frac{1}{6} R^\nu_{\ \lambda} u^\lambda + \int_{-\infty}^{\tau^-} \nabla^\nu G_+\bigl(z(\tau), z(\tau')\bigr)\, d\tau' \biggr] \tag{17.47}$$

for the scalar charge, with $m := m_0 - q\Phi(z)$ denoting the (also formally divergent) dynamical mass of the particle. We see that m and δm combine in Eq. (17.47) to form the particle's observed mass $m_{\rm obs}$, which is taken to be finite and to give a true measure of the particle's inertia. All diverging quantities have thus disappeared into the process of mass renormalization. Substituting Eqs. (17.44) and (17.46) into Eq. (17.8), in which we replace m by $m_{\rm obs} = m + \delta m$, returns an expression for the rate of change of the observed mass,

$$\frac{dm_{\rm obs}}{d\tau} = -\frac{1}{12}(1 - 6\xi) q^2 R - q^2 u^\mu \int_{-\infty}^{\tau^-} \nabla_\mu G_+\bigl(z(\tau), z(\tau')\bigr)\, d\tau'. \tag{17.48}$$

That the observed mass is *not* conserved is a remarkable property of the dynamics of a scalar charge in a curved spacetime. Physically, this corresponds to the fact that in a spacetime with a time-dependent metric, a scalar charge radiates monopole waves and the radiated energy comes at the expense of the particle's inertial mass.

We must confess that the derivation of the equations of motion outlined above returns the *wrong expression* for the self-energy of a spherical shell of scalar charge. We obtained $\delta m = q^2/(3s_0)$, while the correct expression is $\delta m = q^2/(2s_0)$; we are wrong by a factor of 2/3. We believe that this discrepancy originates in a previously stated assumption, that the field on the shell (as produced by the shell itself) is equal to the field of a point particle evaluated at $s = s_0$. We believe that this assumption is in fact wrong, and that a calculation of the field actually produced by a spherical shell would return the correct expression for δm. We also believe, however, that except for the diverging terms that determine δm, the difference between the shell's field and the particle's field should vanish in the limit $s_0 \to 0$. Our conclusion is therefore that while our expression for δm is admittedly incorrect, the statement of the equations of motion is reliable.

Apart from the term proportional to δm, the averaged field of Eq. (17.44) has exactly the same form as the regular field of Eq. (17.42), which we re-express as

$$q\Phi^{\rm R}_\mu = -\frac{1}{12}(1 - 6\xi) q^2 R u_\mu + q^2 \bigl(g_{\mu\nu} + u_\mu u_\nu\bigr) \biggl(\frac{1}{3}\dot{a}^\nu + \frac{1}{6} R^\nu_{\ \lambda} u^\lambda \biggr) + q\Phi^{\rm tail}_\mu. \tag{17.49}$$

The force acting on the point particle can therefore be thought of as originating from the regular field, while the singular field simply contributes to the particle's inertia. After mass renormalization, Eqs. (17.47) and (17.48) are equivalent to the statements

$$m a^\mu = q\bigl(g^{\mu\nu} + u^\mu u^\nu\bigr) \Phi^{\rm R}_\nu(z), \qquad \frac{dm}{d\tau} = -q u^\mu \Phi^{\rm R}_\mu(z), \tag{17.50}$$

where we have dropped the superfluous label "obs" on the particle's observed mass. Another argument in support of the claim that the motion of the particle should be affected by the regular field only was presented in Sec. 14.5.

The equations of motion displayed in Eqs. (17.47) and (17.48) are third-order differential equations for the functions $z^\mu(\tau)$. It is well known that such a system of equations admits many unphysical solutions, such as runaway situations in which the particle's acceleration increases exponentially with τ, even in the

absence of any external force [1,9]. And indeed, our equations of motion do not yet incorporate an external force which presumably is mostly responsible for the particle's acceleration. Both defects can be cured in one stroke. We shall take the point of view, the only admissible one in a classical treatment, that a point particle is merely an idealization for an extended object whose internal structure — the details of its charge distribution — can be considered to be irrelevant. This view automatically implies that our equations are meant to provide only an *approximate* description of the object's motion. It can then be shown [11,12] that within the context of this approximation, it is consistent to replace, on the right-hand side of the equations of motion, any occurrence of the acceleration vector by $f^\mu_{\rm ext}/m$, where $f^\mu_{\rm ext}$ is the external force acting on the particle. Because $f^\mu_{\rm ext}$ is a prescribed quantity, differentiation of the external force does not produce higher derivatives of the functions $z^\mu(\tau)$, and the equations of motion are properly of the second order.

We shall strengthen this conclusion in part V of the review, when we consider the motion of an extended body in a curved external spacetime. While the discussion there will concern the gravitational self-force, many of the lessons learned in part V apply just as well to the case of a scalar (or electric) charge. And the main lesson is this: It is natural — indeed it is an imperative — to view an equation of motion such as Eq. (17.47) as an expansion of the acceleration in powers of q^2, and it is therefore appropriate — indeed imperative — to insert the zeroth-order expression for $\dot{a}^\nu$ within the term of order q^2. The resulting expression for the acceleration is then valid up to correction terms of order q^4. Omitting these error terms, we shall write, in final analysis, the equations of motion in the form

$$m\frac{Du^\mu}{d\tau} = f^\mu_{\rm ext} + q^2\bigl(\delta^\mu_{\ \nu} + u^\mu u_\nu\bigr)\biggl[\frac{1}{3m}\frac{Df^\nu_{\rm ext}}{d\tau} + \frac{1}{6}R^\nu_{\ \lambda}u^\lambda + \int_{-\infty}^{\tau^-}\nabla^\nu G_+\bigl(z(\tau),z(\tau')\bigr)\,d\tau'\biggr] \tag{17.51}$$

and

$$\frac{dm}{d\tau} = -\frac{1}{12}(1-6\xi)q^2R - q^2u^\mu\int_{-\infty}^{\tau^-}\nabla_\mu G_+\bigl(z(\tau),z(\tau')\bigr)\,d\tau', \tag{17.52}$$

where m denotes the observed inertial mass of the scalar charge, and where all tensors are evaluated at $z(\tau)$. We recall that the tail integration must be cut short at $\tau' = \tau^- := \tau - 0^+$ to avoid the singular behaviour of the retarded Green's function at coincidence; this procedure was justified at the beginning of Sec. 17.3. Equations (17.51) and (17.52) were first derived by Theodore C. Quinn in 2000 [8]. In his paper Quinn also establishes that the total work done by the scalar self-force matches the amount of energy radiated away by the particle.

18 Motion of an electric charge

18.1 Dynamics of a point electric charge

A point particle carries an electric charge e and moves on a world line γ described by relations $z^\mu(\lambda)$, in which λ is an arbitrary parameter. The particle generates a vector potential $A^\alpha(x)$ and an electromagnetic field $F_{\alpha\beta}(x) = \nabla_\alpha A_\beta - \nabla_\beta A_\alpha$. The dynamics of the entire system is governed by the action

$$S = S_{\rm field} + S_{\rm particle} + S_{\rm interaction}, \tag{18.1}$$

where $S_{\rm field}$ is an action functional for a free electromagnetic field in a spacetime with metric $g_{\alpha\beta}$, $S_{\rm particle}$ is the action of a free particle moving on a world line γ in this spacetime, and $S_{\rm interaction}$ is an interaction term that couples the field to the particle.

The field action is given by

$$S_{\rm field} = -\frac{1}{16\pi}\int F_{\alpha\beta}F^{\alpha\beta}\sqrt{-g}\,d^4x, \tag{18.2}$$

where the integration is over all of spacetime. The particle action is

$$S_{\rm particle} = -m\int_\gamma d\tau, \tag{18.3}$$

where m is the bare mass of the particle and $d\tau = \sqrt{-g_{\mu\nu}(z)\dot{z}^\mu \dot{z}^\nu}\, d\lambda$ is the differential of proper time along the world line; we use an overdot to indicate differentiation with respect to the parameter λ. Finally, the interaction term is given by

$$S_{\text{interaction}} = e\int_\gamma A_\mu(z)\dot{z}^\mu\, d\lambda = e\int A_\alpha(x) g^\alpha_{\ \mu}(x,z)\dot{z}^\mu \delta_4(x,z)\sqrt{-g}\, d^4x d\lambda. \tag{18.4}$$

Notice that both S_{particle} and $S_{\text{interaction}}$ are invariant under a reparameterization $\lambda \to \lambda'(\lambda)$ of the world line.

Demanding that the total action be stationary under a variation $\delta A^\alpha(x)$ of the vector potential yields Maxwell's equations

$$F^{\alpha\beta}_{\ \ \ ;\beta} = 4\pi j^\alpha \tag{18.5}$$

with a current density $j^\alpha(x)$ defined by

$$j^\alpha(x) = e\int_\gamma g^\alpha_{\ \mu}(x,z)\dot{z}^\mu \delta_4(x,z)\, d\lambda. \tag{18.6}$$

These equations determine the electromagnetic field $F_{\alpha\beta}$ once the motion of the electric charge is specified. On the other hand, demanding that the total action be stationary under a variation $\delta z^\mu(\lambda)$ of the world line yields the equations of motion

$$m\frac{Du^\mu}{d\tau} = eF^\mu_{\ \nu}(z)u^\nu \tag{18.7}$$

for the electric charge. We have adopted τ as the parameter on the world line, and introduced the four-velocity $u^\mu(\tau) := dz^\mu/d\tau$.

The electromagnetic field $F_{\alpha\beta}$ is invariant under a gauge transformation of the form $A_\alpha \to A_\alpha + \nabla_\alpha \Lambda$, in which $\Lambda(x)$ is an arbitrary scalar function. This function can always be chosen so that the vector potential satisfies the Lorenz gauge condition,

$$\nabla_\alpha A^\alpha = 0. \tag{18.8}$$

Under this condition the Maxwell equations of Eq. (18.5) reduce to a wave equation for the vector potential,

$$\Box A^\alpha - R^\alpha_{\ \beta} A^\beta = -4\pi j^\alpha, \tag{18.9}$$

where $\Box = g^{\alpha\beta}\nabla_\alpha\nabla_\beta$ is the wave operator and $R^\alpha_{\ \beta}$ is the Ricci tensor. Having adopted τ as the parameter on the world line, we can re-express the current density of Eq. (18.6) as

$$j^\alpha(x) = e\int_\gamma g^\alpha_{\ \mu}(x,z)u^\mu \delta_4(x,z)\, d\tau, \tag{18.10}$$

and we shall use Eqs. (18.9) and (18.10) to determine the electromagnetic field of a point electric charge. The motion of the particle is in principle determined by Eq. (18.7), but because the vector potential obtained from Eq. (18.9) is singular on the world line, these equations have only formal validity. Before we can make sense of them we will have to analyze the field's singularity structure near the world line. The calculations to be carried out parallel closely those presented in Sec. 17 for the case of a scalar charge; the details will therefore be kept to a minimum and the reader is referred to Sec. 17 for additional information.

18.2 Retarded potential near the world line

The retarded solution to Eq. (18.9) is $A^\alpha(x) = \int G^{\ \alpha}_{+\beta'}(x,x') j^{\beta'}(x')\sqrt{g'}\, d^4x'$, where $G^{\ \alpha}_{+\beta'}(x,x')$ is the retarded Green's function introduced in Sec. 15. After substitution of Eq. (18.10) we obtain

$$A^\alpha(x) = e\int_\gamma G^{\ \alpha}_{+\mu}(x,z)u^\mu\, d\tau, \tag{18.11}$$

in which $z^\mu(\tau)$ gives the description of the world line γ and $u^\mu(\tau) = dz^\mu/d\tau$. Because the retarded Green's function is defined globally in the entire spacetime, Eq. (18.11) applies to any field point x.

We now specialize Eq. (18.11) to a point x close to the world line. We let $\mathcal{N}(x)$ be the normal convex neighbourhood of this point, and we assume that the world line traverses $\mathcal{N}(x)$; refer back to Fig. 9. As in Sec. 17.2 we let $\tau_<$ and $\tau_>$ be the values of the proper-time parameter at which γ enters and leaves $\mathcal{N}(x)$, respectively. Then Eq. (18.11) can be expressed as

$$A^\alpha(x) = e\int_{-\infty}^{\tau_<} G_{+\mu}^{\ \alpha}(x,z)u^\mu\, d\tau + e\int_{\tau_<}^{\tau_>} G_{+\mu}^{\ \alpha}(x,z)u^\mu\, d\tau + e\int_{\tau_>}^{\infty} G_{+\mu}^{\ \alpha}(x,z)u^\mu\, d\tau.$$

The third integration vanishes because x is then in the past of $z(\tau)$, and $G_{+\mu}^{\ \alpha}(x,z) = 0$. For the second integration, x is the normal convex neighbourhood of $z(\tau)$, and the retarded Green's function can be expressed in the Hadamard form produced in Sec. 15.2. This gives

$$\int_{\tau_<}^{\tau_>} G_{+\mu}^{\ \alpha}(x,z)u^\mu\, d\tau = \int_{\tau_<}^{\tau_>} U^\alpha_{\ \mu}(x,z)u^\mu\delta_+(\sigma)\, d\tau + \int_{\tau_<}^{\tau_>} V^\alpha_{\ \mu}(x,z)u^\mu\theta_+(-\sigma)\, d\tau,$$

and to evaluate this we let $x' := z(u)$ be the retarded point associated with x; these points are related by $\sigma(x,x') = 0$ and $r := \sigma_{\alpha'}u^{\alpha'}$ is the retarded distance between x and the world line. To perform the first integration we change variables from τ to σ, noticing that σ increases as $z(\tau)$ passes through x'; the integral evaluates to $U^\alpha_{\ \beta'}u^{\beta'}/r$. The second integration is cut off at $\tau = u$ by the step function, and we obtain our final expression for the vector potential of a point electric charge:

$$A^\alpha(x) = \frac{e}{r}U^\alpha_{\ \beta'}(x,x')u^{\beta'} + e\int_{\tau_<}^{u} V^\alpha_{\ \mu}(x,z)u^\mu\, d\tau + e\int_{-\infty}^{\tau_<} G_{+\mu}^{\ \alpha}(x,z)u^\mu\, d\tau. \tag{18.12}$$

This expression applies to a point x sufficiently close to the world line that there exists a nonempty intersection between $\mathcal{N}(x)$ and γ.

18.3 Electromagnetic field in retarded coordinates

When we differentiate the vector potential of Eq. (18.12) we must keep in mind that a variation in x induces a variation in x', because the new points $x+\delta x$ and $x'+\delta x'$ must also be linked by a null geodesic. Taking this into account, we find that the gradient of the vector potential is given by

$$\nabla_\beta A_\alpha(x) = -\frac{e}{r^2}U_{\alpha\beta'}u^{\beta'}\partial_\beta r + \frac{e}{r}U_{\alpha\beta';\beta}u^{\beta'} + \frac{e}{r}\Bigl(U_{\alpha\beta';\gamma'}u^{\beta'}u^{\gamma'} + U_{\alpha\beta'}a^{\beta'}\Bigr)\partial_\beta u + eV_{\alpha\beta'}u^{\beta'}\partial_\beta u + A^{\rm tail}_{\alpha\beta}(x), \tag{18.13}$$

where the "tail integral" is defined by

$$\begin{aligned} A^{\rm tail}_{\alpha\beta}(x) &= e\int_{\tau_<}^{u} \nabla_\beta V_{\alpha\mu}(x,z)u^\mu\, d\tau + e\int_{-\infty}^{\tau_<} \nabla_\beta G_{+\alpha\mu}(x,z)u^\mu\, d\tau \\ &= e\int_{-\infty}^{u^-} \nabla_\beta G_{+\alpha\mu}(x,z)u^\mu\, d\tau. \end{aligned} \tag{18.14}$$

The second form of the definition, in which we integrate the gradient of the retarded Green's function from $\tau = -\infty$ to $\tau = u^- := u - 0^+$ to avoid the singular behaviour of the retarded Green's function at $\sigma = 0$, is equivalent to the first form.

We shall now expand $F_{\alpha\beta} = \nabla_\alpha A_\beta - \nabla_\beta A_\alpha$ in powers of r, and express the result in terms of the retarded coordinates (u, r, Ω^a) introduced in Sec. 10. It will be convenient to decompose the electromagnetic field in the tetrad (e^α_0, e^α_a) that is obtained by parallel transport of $(u^{\alpha'}, e^{\alpha'}_a)$ on the null geodesic that links x to $x' := z(u)$; this construction is detailed in Sec. 10. We recall from Eq. (10.4) that the parallel propagator can be expressed as $g^{\alpha'}_{\ \alpha} = u^{\alpha'}e^0_\alpha + e^{\alpha'}_a e^a_\alpha$. The expansion relies on Eq. (10.29) for $\partial_\alpha u$, Eq. (10.31) for $\partial_\alpha r$, and we shall need

$$U_{\alpha\beta'}u^{\beta'} = g^{\alpha'}_{\ \alpha}\Bigl[u_{\alpha'} + \frac{1}{12}r^2\bigl(R_{00} + 2R_{0a}\Omega^a + R_{ab}\Omega^a\Omega^b\bigr)u_{\alpha'} + O(r^3)\Bigr], \tag{18.15}$$

which follows from Eq. (15.10) and the relation $\sigma^{\alpha'} = -r(u^{\alpha'} + \Omega^a e^{\alpha'}_a)$ first encountered in Eq. (10.7). We shall also need the expansions

$$U_{\alpha\beta';\beta} u^{\beta'} = -\frac{1}{2} r g^{\alpha'}_{\ \alpha} g^{\beta'}_{\ \beta} \Big[R_{\alpha' 0\beta' 0} + R_{\alpha' 0 \beta' c}\Omega^c - \frac{1}{3}\big(R_{\beta' 0} + R_{\beta' c}\Omega^c\big) u_{\alpha'} + O(r) \Big] \tag{18.16}$$

and

$$U_{\alpha\beta';\gamma'} u^{\beta'} u^{\gamma'} + U_{\alpha\beta'} a^{\beta'} = g^{\alpha'}_{\ \alpha} \Big[a_{\alpha'} + \frac{1}{2} r R_{\alpha' 0 b 0}\Omega^b - \frac{1}{6} r \big(R_{00} + R_{0b}\Omega^b\big) u_{\alpha'} + O(r^2) \Big] \tag{18.17}$$

that follow from Eqs. (15.10)–(15.12). And finally, we shall need

$$V_{\alpha\beta'} u^{\beta'} = -\frac{1}{2} g^{\alpha'}_{\ \alpha} \Big[R_{\alpha' 0} - \frac{1}{6} R u_{\alpha'} + O(r) \Big], \tag{18.18}$$

a relation that was first established in Eq. (15.14).

Collecting all these results gives

$$\begin{aligned}
F_{a0}(u,r,\Omega^a) &:= F_{\alpha\beta}(x) e^\alpha_a(x) e^\beta_0(x) \\
&= \frac{e}{r^2}\Omega_a - \frac{e}{r}\big(a_a - a_b\Omega^b\Omega_a\big) + \frac{1}{3} e R_{b0c0}\Omega^b\Omega^c\Omega_a - \frac{1}{6} e \big(5R_{a0b0}\Omega^b + R_{ab0c}\Omega^b\Omega^c\big) \\
&\quad + \frac{1}{12} e\big(5R_{00} + R_{bc}\Omega^b\Omega^c + R\big)\Omega_a + \frac{1}{3} e R_{a0} - \frac{1}{6} e R_{ab}\Omega^b + F^{\rm tail}_{a0} + O(r), \qquad (18.19) \\
F_{ab}(u,r,\Omega^a) &:= F_{\alpha\beta}(x) e^\alpha_a(x) e^\beta_b(x) \\
&= \frac{e}{r}\big(a_a\Omega_b - \Omega_a a_b\big) + \frac{1}{2} e \big(R_{a0bc} - R_{b0ac} + R_{a0c0}\Omega_b - \Omega_a R_{b0c0}\big)\Omega^c \\
&\quad - \frac{1}{2} e\big(R_{a0}\Omega_b - \Omega_a R_{b0}\big) + F^{\rm tail}_{ab} + O(r), \qquad (18.20)
\end{aligned}$$

where

$$F^{\rm tail}_{a0} = F^{\rm tail}_{\alpha'\beta'}(x') e^{\alpha'}_a u^{\beta'}, \qquad F^{\rm tail}_{ab} = F^{\rm tail}_{\alpha'\beta'}(x') e^{\alpha'}_a e^{\beta'}_b \tag{18.21}$$

are the frame components of the tail integral; this is obtained from Eq. (18.14) evaluated at x':

$$F^{\rm tail}_{\alpha'\beta'}(x') = 2e \int_{-\infty}^{u^-} \nabla_{[\alpha'} G_{+\beta']\mu}(x', z) u^\mu \, d\tau. \tag{18.22}$$

It should be emphasized that in Eqs. (18.19) and (18.20), all frame components are evaluated at the retarded point $x' := z(u)$ associated with x; for example, $a_a := a_a(u) := a_{\alpha'} e^{\alpha'}_a$. It is clear from these equations that the electromagnetic field $F_{\alpha\beta}(x)$ is singular on the world line.

18.4 Electromagnetic field in Fermi normal coordinates

We now wish to express the electromagnetic field in the Fermi normal coordinates of Sec. 9; as before those will be denoted (t, s, ω^a). The translation will be carried out as in Sec. 17.4, and we will decompose the field in the tetrad $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ that is obtained by parallel transport of $(u^{\bar{\alpha}}, e^{\bar{\alpha}}_a)$ on the spacelike geodesic that links x to the simultaneous point $\bar{x} := z(t)$.

Our first task is to decompose $F_{\alpha\beta}(x)$ in the tetrad $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$, thereby defining $\bar{F}_{a0} := F_{\alpha\beta}\bar{e}^\alpha_a\bar{e}^\beta_0$ and $\bar{F}_{ab} := F_{\alpha\beta}\bar{e}^\alpha_a\bar{e}^\beta_b$. For this purpose we use Eqs. (11.7), (11.8), (18.19), and (18.20) to obtain

$$\begin{aligned}
\bar{F}_{a0} &= \frac{e}{r^2}\Omega_a - \frac{e}{r}\big(a_a - a_b\Omega^b\Omega_a\big) + \frac{1}{2} e a_b\Omega^b a_a + \frac{1}{2} e\dot{a}_0\Omega_a - \frac{5}{6} e R_{a0b0}\Omega^b + \frac{1}{3} e R_{b0c0}\Omega^b\Omega^c\Omega_a \\
&\quad + \frac{1}{3} e R_{ab0c}\Omega^b\Omega^c + \frac{1}{12} e \big(5R_{00} + R_{bc}\Omega^b\Omega^c + R\big)\Omega_a + \frac{1}{3} e R_{a0} - \frac{1}{6} e R_{ab}\Omega^b + \bar{F}^{\rm tail}_{a0} + O(r)
\end{aligned}$$

and

$$\bar{F}_{ab} = \frac{1}{2} e\big(\Omega_a\dot{a}_b - \dot{a}_a\Omega_b\big) + \frac{1}{2} e\big(R_{a0bc} - R_{b0ac}\big)\Omega^c - \frac{1}{2} e\big(R_{a0}\Omega_b - \Omega_a R_{b0}\big) + \bar{F}^{\rm tail}_{ab} + O(r),$$

where all frame components are still evaluated at x', except for

$$\bar{F}^{\rm tail}_{a0} := F^{\rm tail}_{\bar\alpha\bar\beta}(\bar x)e^{\bar\alpha}_a u^{\bar\beta}, \qquad \bar{F}^{\rm tail}_{ab} := F^{\rm tail}_{\bar\alpha\bar\beta}(\bar x)e^{\bar\alpha}_a e^{\bar\beta}_b,$$

which are evaluated at $\bar x$.

We must still translate these results into the Fermi normal coordinates (t, s, ω^a). For this we involve Eqs. (11.4), (11.5), and (11.6), and we recycle some computations that were first carried out in Sec. 17.4. After some algebra, we arrive at

$$\begin{aligned}
\bar F_{a0}(t,s,\omega^a) &:= F_{\alpha\beta}(x)\bar e^\alpha_a(x)\bar e^\beta_0(x) \\
&= \frac{e}{s^2}\omega_a - \frac{e}{2s}\bigl(a_a + a_b\omega^b\omega_a\bigr) + \frac{3}{4}ea_b\omega^b a_a + \frac{3}{8}e\bigl(a_b\omega^b\bigr)^2\omega_a + \frac{3}{8}e\dot a_0\omega_a + \frac{2}{3}e\dot a_a \\
&\quad - \frac{2}{3}eR_{a0b0}\omega^b - \frac{1}{6}eR_{b0c0}\omega^b\omega^c\omega_a + \frac{1}{12}e\bigl(5R_{00} + R_{bc}\omega^b\omega^c + R\bigr)\omega_a \\
&\quad + \frac{1}{3}eR_{a0} - \frac{1}{6}eR_{ab}\omega^b + \bar F^{\rm tail}_{a0} + O(s), \qquad (18.23)\\
\bar F_{ab}(t,s,\omega^a) &:= F_{\alpha\beta}(x)\bar e^\alpha_a(x)\bar e^\beta_b(x) \\
&= \frac{1}{2}e\bigl(\omega_a\dot a_b - \dot a_a\omega_b\bigr) + \frac{1}{2}e\bigl(R_{a0bc} - R_{b0ac}\bigr)\omega^c - \frac{1}{2}e\bigl(R_{a0}\omega_b - \omega_a R_{b0}\bigr) \\
&\quad + \bar F^{\rm tail}_{ab} + O(s), \qquad (18.24)
\end{aligned}$$

where all frame components are now evaluated at $\bar x := z(t)$; for example, $a_a := a_a(t) := a_{\bar\alpha}e^{\bar\alpha}_a$.

Our next task is to compute the averages of $\bar F_{a0}$ and $\bar F_{ab}$ over $S(t,s)$, a two-surface of constant t and s. These are defined by

$$\bigl\langle \bar F_{a0}\bigr\rangle(t,s) = \frac{1}{\mathcal{A}}\oint_{S(t,s)} \bar F_{a0}(t,s,\omega^a)\,d\mathcal{A}, \qquad \bigl\langle \bar F_{ab}\bigr\rangle(t,s) = \frac{1}{\mathcal{A}}\oint_{S(t,s)} \bar F_{ab}(t,s,\omega^a)\,d\mathcal{A}, \qquad (18.25)$$

where $d\mathcal{A}$ is the element of surface area on $S(t,s)$, and $\mathcal{A} = \oint d\mathcal{A}$. Using the methods developed in Sec. 17.4, we find

$$\begin{aligned}
\bigl\langle \bar F_{a0}\bigr\rangle &= -\frac{2e}{3s}a_a + \frac{2}{3}e\dot a_a + \frac{1}{3}eR_{a0} + \bar F^{\rm tail}_{a0} + O(s), \qquad (18.26)\\
\bigl\langle \bar F_{ab}\bigr\rangle &= \bar F^{\rm tail}_{ab} + O(s). \qquad (18.27)
\end{aligned}$$

The averaged field is singular on the world line, but we nevertheless take the formal limit $s \to 0$ of the expressions displayed in Eqs. (18.26) and (18.27). In the limit the tetrad $(\bar e^\alpha_0, \bar e^\alpha_a)$ becomes $(u^{\bar\alpha}, e^{\bar\alpha}_a)$, and we can easily reconstruct the field at $\bar x$ from its frame components. We thus obtain

$$\bigl\langle F_{\bar\alpha\bar\beta}\bigr\rangle = \lim_{s\to 0}\biggl(-\frac{4e}{3s}\biggr)u_{[\bar\alpha}a_{\bar\beta]} + 2eu_{[\bar\alpha}\bigl(g_{\bar\beta]\bar\gamma} + u_{\bar\beta]}u_{\bar\gamma}\bigr)\biggl(\frac{2}{3}\dot a^{\bar\gamma} + \frac{1}{3}R^{\bar\gamma}_{\ \bar\delta}u^{\bar\delta}\biggr) + F^{\rm tail}_{\bar\alpha\bar\beta}, \qquad (18.28)$$

where the tail term can be copied from Eq. (18.22),

$$F^{\rm tail}_{\bar\alpha\bar\beta}(\bar x) = 2e\int_{-\infty}^{t^-} \nabla_{[\bar\alpha}G_{+\bar\beta]\mu}(\bar x, z)u^\mu\,d\tau. \qquad (18.29)$$

The tensors appearing in Eq. (18.28) all refer to $\bar x := z(t)$, which now stands for an arbitrary point on the world line γ.

18.5 Singular and regular fields

The singular vector potential

$$A^\alpha_{\rm S}(x) = e\int_\gamma G_{{\rm S}\ \mu}^{\ \alpha}(x,z)u^\mu\,d\tau \qquad (18.30)$$

is the (unphysical) solution to Eqs. (18.9) and (18.10) that is obtained by adopting the singular Green's function of Eq. (15.24) instead of the retarded Green's function. We will see that the singular field $F^{\rm S}_{\alpha\beta}$ reproduces the singular behaviour of the retarded solution, and that the difference, $F^{\rm R}_{\alpha\beta} = F_{\alpha\beta} - F^{\rm S}_{\alpha\beta}$, is regular on the world line.

To evaluate the integral of Eq. (18.30) we assume once more that x is sufficiently close to γ that the world line traverses $\mathcal{N}(x)$; refer back to Fig. 9. As before we let $\tau_<$ and $\tau_>$ be the values of the proper-time parameter at which γ enters and leaves $\mathcal{N}(x)$, respectively. Then Eq. (18.30) becomes

$$A^\alpha_{\rm S}(x) = e\int_{-\infty}^{\tau_<} G^{\ \alpha}_{{\rm S}\ \mu}(x,z)u^\mu\, d\tau + e\int_{\tau_<}^{\tau_>} G^{\ \alpha}_{{\rm S}\ \mu}(x,z)u^\mu\, d\tau + e\int_{\tau_>}^{\infty} G^{\ \alpha}_{{\rm S}\ \mu}(x,z)u^\mu\, d\tau.$$

The first integration vanishes because x is then in the chronological future of $z(\tau)$, and $G^{\ \alpha}_{{\rm S}\ \mu}(x,z) = 0$ by Eq. (15.27). Similarly, the third integration vanishes because x is then in the chronological past of $z(\tau)$. For the second integration, x is the normal convex neighbourhood of $z(\tau)$, the singular Green's function can be expressed in the Hadamard form of Eq. (15.33), and we have

$$\begin{aligned}\int_{\tau_<}^{\tau_>} G^{\ \alpha}_{{\rm S}\ \mu}(x,z)u^\mu\, d\tau &= \frac{1}{2}\int_{\tau_<}^{\tau_>} U^\alpha_{\ \mu}(x,z)u^\mu\delta_+(\sigma)\, d\tau + \frac{1}{2}\int_{\tau_<}^{\tau_>} U^\alpha_{\ \mu}(x,z)u^\mu\delta_-(\sigma)\, d\tau \\ &\quad - \frac{1}{2}\int_{\tau_<}^{\tau_>} V^\alpha_{\ \mu}(x,z)u^\mu\theta(\sigma)\, d\tau.\end{aligned}$$

To evaluate these we let $x' := z(u)$ and $x'' := z(v)$ be the retarded and advanced points associated with x, respectively. To perform the first integration we change variables from τ to σ, noticing that σ increases as $z(\tau)$ passes through x'; the integral evaluates to $U^\alpha_{\ \beta'}u^{\beta'}/r$. We do the same for the second integration, but we notice now that σ decreases as $z(\tau)$ passes through x''; the integral evaluates to $U^\alpha_{\ \beta''}u^{\beta''}/r_{\rm adv}$, where $r_{\rm adv} := -\sigma_{\alpha''}u^{\alpha''}$ is the advanced distance between x and the world line. The third integration is restricted to the interval $u \le \tau \le v$ by the step function, and we obtain the expression

$$A^\alpha_{\rm S}(x) = \frac{e}{2r}U^\alpha_{\ \beta'}u^{\beta'} + \frac{e}{2r_{\rm adv}}U^\alpha_{\ \beta''}u^{\beta''} - \frac{1}{2}e\int_u^v V^\alpha_{\ \mu}(x,z)u^\mu\, d\tau \tag{18.31}$$

for the singular vector potential.

Differentiation of Eq. (18.31) yields

$$\begin{aligned}\nabla_\beta A^{\rm S}_\alpha(x) &= -\frac{e}{2r^2}U_{\alpha\beta'}u^{\beta'}\partial_\beta r - \frac{e}{2{r_{\rm adv}}^2}U_{\alpha\beta''}u^{\beta''}\partial_\beta r_{\rm adv} + \frac{e}{2r}U_{\alpha\beta';\beta}u^{\beta'} \\ &\quad + \frac{e}{2r}\Bigl(U_{\alpha\beta';\gamma'}u^{\beta'}u^{\gamma'} + U_{\alpha\beta'}a^{\beta'}\Bigr)\partial_\beta u + \frac{e}{2r_{\rm adv}}U_{\alpha\beta'';\beta}u^{\beta''} \\ &\quad + \frac{e}{2r_{\rm adv}}\Bigl(U_{\alpha\beta'';\gamma''}u^{\beta''}u^{\gamma''} + U_{\alpha\beta''}a^{\beta''}\Bigr)\partial_\beta v + \frac{1}{2}eV_{\alpha\beta'}u^{\beta'}\partial_\beta u \\ &\quad - \frac{1}{2}eV_{\alpha\beta''}u^{\beta''}\partial_\beta v - \frac{1}{2}e\int_u^v \nabla_\beta V_{\alpha\mu}(x,z)u^\mu\, d\tau,\end{aligned} \tag{18.32}$$

and we would like to express this as an expansion in powers of r. For this we will rely on results already established in Sec. 18.3, as well as additional expansions that will involve the advanced point x''. We recall that a relation between retarded and advanced times was worked out in Eq. (11.12), that an expression for the advanced distance was displayed in Eq. (11.13), and that Eqs. (11.14) and (11.15) give expansions for $\partial_\alpha v$ and $\partial_\alpha r_{\rm adv}$, respectively.

To derive an expansion for $U_{\alpha\beta''}u^{\beta''}$ we follow the general method of Sec. 11.4 and introduce the functions $U_\alpha(\tau) := U_{\alpha\mu}(x,z)u^\mu$. We have that

$$U_{\alpha\beta''}u^{\beta''} := U_\alpha(v) = U_\alpha(u) + \dot{U}_\alpha(u)\Delta' + \frac{1}{2}\ddot{U}_\alpha(u)\Delta'^2 + O\bigl(\Delta'^3\bigr),$$

where overdots indicate differentiation with respect to τ, and $\Delta' := v - u$. The leading term $U_\alpha(u) := U_{\alpha\beta'}u^{\beta'}$ was worked out in Eq. (18.15), and the derivatives of $U_\alpha(\tau)$ are given by

$$\dot{U}_\alpha(u) = U_{\alpha\beta';\gamma'}u^{\beta'}u^{\gamma'} + U_{\alpha\beta'}a^{\beta'} = g^{\alpha'}_{\ \alpha}\Bigl[a_{\alpha'} + \frac{1}{2}rR_{\alpha'0b0}\Omega^b - \frac{1}{6}r\bigl(R_{00} + R_{0b}\Omega^b\bigr)u_{\alpha'} + O(r^2)\Bigr]$$

and

$$\ddot{U}_\alpha(u) = U_{\alpha\beta';\gamma'\delta'}u^{\beta'}u^{\gamma'}u^{\delta'} + U_{\alpha\beta';\gamma'}\big(2a^{\beta'}u^{\gamma'} + u^{\beta'}a^{\gamma'}\big) + U_{\alpha\beta'}\dot{a}^{\beta'} = g^{\alpha'}_{\ \alpha}\Big[\dot{a}_{\alpha'} + \frac{1}{6}R_{00}u_{\alpha'} + O(r)\Big],$$

according to Eqs. (18.17) and (15.12). Combining these results together with Eq. (11.12) for Δ' gives

$$\begin{aligned} U_{\alpha\beta''}u^{\beta''} &= g^{\alpha'}_{\ \alpha}\Big[u_{\alpha'} + 2r\big(1 - ra_b\Omega^b\big)a_{\alpha'} + 2r^2\dot{a}_{\alpha'} + r^2 R_{\alpha'0b0}\Omega^b \\ &\qquad + \frac{1}{12}r^2\big(R_{00} - 2R_{0a}\Omega^a + R_{ab}\Omega^a\Omega^b\big)u_{\alpha'} + O(r^3)\Big], \end{aligned} \tag{18.33}$$

which should be compared with Eq. (18.15). It should be emphasized that in Eq. (18.33) and all equations below, all frame components are evaluated at the retarded point x', and not at the advanced point. The preceding computation gives us also an expansion for

$$U_{\alpha\beta'';\gamma''}u^{\beta''}u^{\gamma''} + U_{\alpha\beta''}a^{\beta''} := \dot{U}_\alpha(v) = \dot{U}_\alpha(u) + \ddot{U}_\alpha(u)\Delta' + O(\Delta'^2),$$

which becomes

$$U_{\alpha\beta'';\gamma''}u^{\beta''}u^{\gamma''} + U_{\alpha\beta''}a^{\beta''} = g^{\alpha'}_{\ \alpha}\Big[a_{\alpha'} + 2r\dot{a}_{\alpha'} + \frac{1}{2}rR_{\alpha'0b0}\Omega^b + \frac{1}{6}r\big(R_{00} - R_{0b}\Omega^b\big)u_{\alpha'} + O(r^2)\Big], \tag{18.34}$$

and which should be compared with Eq. (18.17).

We proceed similarly to derive an expansion for $U_{\alpha\beta'';\beta}u^{\beta''}$. Here we introduce the functions $U_{\alpha\beta}(\tau) := U_{\alpha\mu;\beta}u^\mu$ and express $U_{\alpha\beta'';\beta}u^{\beta''}$ as $U_{\alpha\beta}(v) = U_{\alpha\beta}(u) + \dot{U}_{\alpha\beta}(u)\Delta' + O(\Delta'^2)$. The leading term $U_{\alpha\beta}(u) := U_{\alpha\beta';\beta}u^{\beta'}$ was computed in Eq. (18.16), and

$$\dot{U}_{\alpha\beta}(u) = U_{\alpha\beta';\beta\gamma'}u^{\beta'}u^{\gamma'} + U_{\alpha\beta';\beta}a^{\beta'} = \frac{1}{2}g^{\alpha'}_{\ \alpha}g^{\beta'}_{\ \beta}\Big[R_{\alpha'0\beta'0} - \frac{1}{3}u_{\alpha'}R_{\beta'0} + O(r)\Big]$$

follows from Eq. (15.11). Combining these results together with Eq. (11.12) for Δ' gives

$$U_{\alpha\beta'';\beta}u^{\beta''} = \frac{1}{2}rg^{\alpha'}_{\ \alpha}g^{\beta'}_{\ \beta}\Big[R_{\alpha'0\beta'0} - R_{\alpha'0\beta'c}\Omega^c - \frac{1}{3}\big(R_{\beta'0} - R_{\beta'c}\Omega^c\big)u_{\alpha'} + O(r)\Big], \tag{18.35}$$

and this should be compared with Eq. (18.16). The last expansion we shall need is

$$V_{\alpha\beta''}u^{\beta''} = -\frac{1}{2}g^{\alpha'}_{\ \alpha}\Big[R_{\alpha'0} - \frac{1}{6}Ru_{\alpha'} + O(r)\Big], \tag{18.36}$$

which follows at once from Eq. (18.18).

It is now a straightforward (but still tedious) matter to substitute these expansions into Eq. (18.32) to obtain the projections of the singular electromagnetic field $F^{\rm S}_{\alpha\beta} = \nabla_\alpha A^{\rm S}_\beta - \nabla_\beta A^{\rm S}_\alpha$ in the same tetrad (e^α_0, e^α_a) that was employed in Sec. 18.3. This gives

$$\begin{aligned} F^{\rm S}_{a0}(u,r,\Omega^a) &:= F^{\rm S}_{\alpha\beta}(x)e^\alpha_a(x)e^\beta_0(x) \\ &= \frac{e}{r^2}\Omega_a - \frac{e}{r}\big(a_a - a_b\Omega^b\Omega_a\big) - \frac{2}{3}e\dot{a}_a + \frac{1}{3}eR_{b0c0}\Omega^b\Omega^c\Omega_a - \frac{1}{6}e\big(5R_{a0b0}\Omega^b + R_{ab0c}\Omega^b\Omega^c\big) \\ &\quad + \frac{1}{12}e\big(5R_{00} + R_{bc}\Omega^b\Omega^c + R\big)\Omega_a - \frac{1}{6}eR_{ab}\Omega^b + O(r), \end{aligned} \tag{18.37}$$

$$\begin{aligned} F^{\rm S}_{ab}(u,r,\Omega^a) &:= F^{\rm S}_{\alpha\beta}(x)e^\alpha_a(x)e^\beta_b(x) \\ &= \frac{e}{r}\big(a_a\Omega_b - \Omega_a a_b\big) + \frac{1}{2}e\big(R_{a0bc} - R_{b0ac} + R_{a0c0}\Omega_b - \Omega_a R_{b0c0}\big)\Omega^c \\ &\quad - \frac{1}{2}e\big(R_{a0}\Omega_b - \Omega_a R_{b0}\big) + O(r), \end{aligned} \tag{18.38}$$

in which all frame components are evaluated at the retarded point x'. Comparison of these expressions with Eqs. (18.19) and (18.20) reveals that the retarded and singular fields share the same singularity structure.

The difference between the retarded field of Eqs. (18.19), (18.20) and the singular field of Eqs. (18.37), (18.38) defines the regular field $F^{\rm R}_{\alpha\beta}(x)$. Its tetrad components are

$$F^{\rm R}_{a0} = \frac{2}{3}e\dot{a}_a + \frac{1}{3}eR_{a0} + F^{\rm tail}_{a0} + O(r), \tag{18.39}$$

$$F^{\rm R}_{ab} = F^{\rm tail}_{ab} + O(r), \tag{18.40}$$

and we see that $F^{\rm R}_{\alpha\beta}$ is a regular tensor field on the world line. There is therefore no obstacle in evaluating the regular field directly at $x = x'$, where the tetrad (e^α_0, e^α_a) becomes $(u^{\alpha'}, e^{\alpha'}_a)$. Reconstructing the field at x' from its frame components, we obtain

$$F^{\rm R}_{\alpha'\beta'}(x') = 2eu_{[\alpha'}\bigl(g_{\beta']\gamma'} + u_{\beta']}u_{\gamma'}\bigr)\biggl(\frac{2}{3}\dot{a}^{\gamma'} + \frac{1}{3}R^{\gamma'}_{\ \delta'}u^{\delta'}\biggr) + F^{\rm tail}_{\alpha'\beta'}, \tag{18.41}$$

where the tail term can be copied from Eq. (18.22),

$$F^{\rm tail}_{\alpha'\beta'}(x') = 2e\int_{-\infty}^{u^-} \nabla_{[\alpha'}G_{+\beta']\mu}(x', z)u^\mu\, d\tau. \tag{18.42}$$

The tensors appearing in Eq. (18.41) all refer to the retarded point $x' := z(u)$, which now stands for an arbitrary point on the world line γ.

18.6 Equations of motion

The retarded field $F_{\alpha\beta}$ of a point electric charge is singular on the world line, and this behaviour makes it difficult to understand how the field is supposed to act on the particle and exert a force. The field's singularity structure was analyzed in Secs. 18.3 and 18.4, and in Sec. 18.5 it was shown to originate from the singular field $F^{\rm S}_{\alpha\beta}$; the regular field $F^{\rm R}_{\alpha\beta} = F_{\alpha\beta} - F^{\rm S}_{\alpha\beta}$ was then shown to regular on the world line.

To make sense of the retarded field's action on the particle we follow the discussion of Sec. 17.6 and temporarily picture the electric charge as a spherical hollow shell; the shell's radius is s_0 in Fermi normal coordinates, and it is independent of the angles contained in the unit vector ω^a. The *net force* acting at proper time τ on this shell is proportional to the average of $F_{\alpha\beta}(\tau, s_0, \omega^a)$ over the shell's surface. Assuming that the field on the shell is equal to the field of a point particle evaluated at $s = s_0$, and ignoring terms that disappear in the limit $s_0 \to 0$, we obtain from Eq. (18.28)

$$e\langle F_{\mu\nu}\rangle u^\nu = -(\delta m)a_\mu + e^2\bigl(g_{\mu\nu} + u_\mu u_\nu\bigr)\biggl(\frac{2}{3}\dot{a}^\nu + \frac{1}{3}R^\nu_{\ \lambda}u^\lambda\biggr) + eF^{\rm tail}_{\mu\nu}u^\nu, \tag{18.43}$$

where

$$\delta m := \lim_{s_0\to 0}\frac{2e^2}{3s_0} \tag{18.44}$$

is formally a divergent quantity and

$$eF^{\rm tail}_{\mu\nu}u^\nu = 2e^2u^\nu\int_{-\infty}^{\tau^-}\nabla_{[\mu}G_{+\nu]\lambda'}\bigl(z(\tau), z(\tau')\bigr)u^{\lambda'}\, d\tau' \tag{18.45}$$

is the tail part of the force; all tensors in Eq. (18.43) are evaluated at an arbitrary point $z(\tau)$ on the world line.

Substituting Eqs. (18.43) and (18.45) into Eq. (18.7) gives rise to the equations of motion

$$(m + \delta m)a^\mu = e^2\bigl(\delta^\mu_{\ \nu} + u^\mu u_\nu\bigr)\biggl(\frac{2}{3}\dot{a}^\nu + \frac{1}{3}R^\nu_{\ \lambda}u^\lambda\biggr) + 2e^2u_\nu\int_{-\infty}^{\tau^-}\nabla^{[\mu}G_{+\ \ \lambda'}^{\ \nu]}\bigl(z(\tau), z(\tau')\bigr)u^{\lambda'}\, d\tau' \tag{18.46}$$

for the electric charge, with m denoting the (also formally divergent) bare mass of the particle. We see that m and δm combine in Eq. (18.46) to form the particle's observed mass $m_{\rm obs}$, which is finite and gives a true

measure of the particle's inertia. All diverging quantities have thus disappeared into the procedure of mass renormalization.

We must confess, as we did in the case of the scalar self-force, that the derivation of the equations of motion outlined above returns the *wrong expression* for the self-energy of a spherical shell of electric charge. We obtained $\delta m = 2e^2/(3s_0)$, while the correct expression is $\delta m = e^2/(2s_0)$; we are wrong by a factor of 4/3. As before we believe that this discrepancy originates in a previously stated assumption, that the field on the shell (as produced by the shell itself) is equal to the field of a point particle evaluated at $s = s_0$. We believe that this assumption is in fact wrong, and that a calculation of the field actually produced by a spherical shell would return the correct expression for δm. We also believe, however, that except for the diverging terms that determine δm, the difference between the shell's field and the particle's field should vanish in the limit $s_0 \to 0$. Our conclusion is therefore that while our expression for δm is admittedly incorrect, the statement of the equations of motion is reliable.

Apart from the term proportional to δm, the averaged force of Eq. (18.43) has exactly the same form as the force that arises from the regular field of Eq. (18.41), which we express as

$$eF^{\mathrm{R}}_{\mu\nu}u^\nu = e^2\bigl(g_{\mu\nu} + u_\mu u_\nu\bigr)\biggl(\frac{2}{3}\dot{a}^\nu + \frac{1}{3}R^\nu{}_\lambda u^\lambda\biggr) + eF^{\mathrm{tail}}_{\mu\nu}u^\nu. \tag{18.47}$$

The force acting on the point particle can therefore be thought of as originating from the regular field, while the singular field simply contributes to the particle's inertia. After mass renormalization, Eq. (18.46) is equivalent to the statement

$$ma_\mu = eF^{\mathrm{R}}_{\mu\nu}(z)u^\nu, \tag{18.48}$$

where we have dropped the superfluous label "obs" on the particle's observed mass.

For the final expression of the equations of motion we follow the discussion of Sec. 17.6 and allow an external force f^μ_{ext} to act on the particle, and we replace, on the right-hand side of the equations, the acceleration vector by f^μ_{ext}/m. This produces

$$m\frac{Du^\mu}{d\tau} = f^\mu_{\mathrm{ext}} + e^2\bigl(\delta^\mu{}_\nu + u^\mu u_\nu\bigr)\biggl(\frac{2}{3m}\frac{Df^\nu_{\mathrm{ext}}}{d\tau} + \frac{1}{3}R^\nu{}_\lambda u^\lambda\biggr) + 2e^2u_\nu\int_{-\infty}^{\tau^-}\nabla^{[\mu}G_+{}^{\nu]}{}_{\lambda'}\bigl(z(\tau),z(\tau')\bigr)u^{\lambda'}\,d\tau', \tag{18.49}$$

in which m denotes the observed inertial mass of the electric charge and all tensors are evaluated at $z(\tau)$, the current position of the particle on the world line; the primed indices in the tail integral refer to the point $z(\tau')$, which represents a prior position. We recall that the integration must be cut short at $\tau' = \tau^- := \tau - 0^+$ to avoid the singular behaviour of the retarded Green's function at coincidence; this procedure was justified at the beginning of Sec. 18.3. Equation (18.49) was first derived (without the Ricci-tensor term) by Bryce S. DeWitt and Robert W. Brehme in 1960 [4], and then corrected by J.M. Hobbs in 1968 [5]. An alternative derivation was produced by Theodore C. Quinn and Robert M. Wald in 1997 [7]. In a subsequent publication [141], Quinn and Wald proved that the total work done by the electromagnetic self-force matches the energy radiated away by the particle.

19 Motion of a point mass

19.1 Dynamics of a point mass

Introduction

In this section we consider the motion of a point particle of mass m subjected to its own gravitational field in addition to an external field. The particle moves on a world line γ in a curved spacetime whose background metric $g_{\alpha\beta}$ is assumed to be a *vacuum* solution to the Einstein field equations. We shall suppose that m is small, so that the perturbation $h_{\alpha\beta}$ created by the particle can also be considered to be small. In the final analysis we shall find that $h_{\alpha\beta}$ obeys a linear wave equation in the background spacetime, and this linearization of the field equations will allow us to fit the problem of determining the motion of a point mass within the general framework developed in Secs. 17 and 18. We shall find that γ is not a geodesic of the background spacetime because $h_{\alpha\beta}$ acts on the particle and produces an acceleration proportional to m; the motion is geodesic in the test-mass limit only.

While we can make the problem fit within the general framework, it is important to understand that the problem of motion in gravitation is conceptually very different from the versions encountered previously in the case of a scalar or electromagnetic field. In these cases, the field equations satisfied by the scalar potential Φ or the vector potential A^α are fundamentally linear; in general relativity the field equations satisfied by $h_{\alpha\beta}$ are fundamentally nonlinear, and this makes a major impact on the formulation of the problem. (In all cases the coupled problem of determining the field *and* the motion of the particle is nonlinear.) Another difference resides with the fact that in the previous cases, the field equations and the law of motion could be formulated independently of each other (because the action functional could be varied independently with respect to the field and the world line); in general relativity the law of motion follows from energy-momentum conservation, which is itself a consequence of the field equations.

The dynamics of a point mass in general relativity must therefore be formulated with care. We shall describe a formal approach to this problem, based on the fiction that the spacetime of a point particle can be constructed exactly in general relativity. (This is indeed a fiction, because it is known [142] that the metric of a point particle, as described by a Dirac distribution on a world line, is much too singular to be defined as a distribution in spacetime. The construction, however, makes distributional sense at the level of the linearized theory.) The outcome of this approach will be an approximate formulation of the equations of motion that relies on a linearization of the field equations, and which turns out to be closely analogous to the scalar and electromagnetic cases encountered previously. We shall put the motion of a small mass on a much sounder foundation in Part V, where we take m to be a (small) extended body instead of a point particle.

Exact formulation

Let a point particle of mass m move on a world line γ in a curved spacetime with metric $\mathsf{g}_{\alpha\beta}$. This is the *exact metric* of the *perturbed spacetime*, and it depends on m as well as all other relevant parameters. At a later stage of the discussion $\mathsf{g}_{\alpha\beta}$ will be expressed as sum of a "background" part $g_{\alpha\beta}$ that is independent of m, and a "perturbation" part $h_{\alpha\beta}$ that contains the dependence on m. The world line is described by relations $z^\mu(\lambda)$ in which λ is an arbitrary parameter — this will later be identified with proper time τ in the *background* spacetime. We use sans-serif symbols to denote tensors that refer to the perturbed spacetime; tensors in the background spacetime will be denoted, as usual, by italic symbols.

The particle's action functional is

$$S_{\text{particle}} = -m \int_\gamma \sqrt{-\mathsf{g}_{\mu\nu}\dot{z}^\mu \dot{z}^\nu}\, d\lambda \tag{19.1}$$

where $\dot{z}^\mu = dz^\mu/d\lambda$ is tangent to the world line and the metric is evaluated at z. We assume that the particle provides the *only* source of matter in the spacetime — an explanation will be provided below — so that the Einstein field equations take the form of

$$\mathsf{G}^{\alpha\beta} = 8\pi \mathsf{T}^{\alpha\beta}, \tag{19.2}$$

where $\mathsf{G}^{\alpha\beta}$ is the Einstein tensor constructed from $\mathsf{g}_{\alpha\beta}$ and

$$\mathsf{T}^{\alpha\beta}(x) = m \int_\gamma \frac{\mathsf{g}^\alpha_{\ \mu}(x,z)\mathsf{g}^\beta_{\ \nu}(x,z)\dot{z}^\mu \dot{z}^\nu}{\sqrt{-\mathsf{g}_{\mu\nu}\dot{z}^\mu \dot{z}^\nu}}\, \delta_4(x,z)\, d\lambda \tag{19.3}$$

is the particle's energy-momentum tensor, obtained by functional differentiation of S_{particle} with respect to $\mathsf{g}_{\alpha\beta}(x)$; the parallel propagators appear naturally by expressing $\mathsf{g}_{\mu\nu}$ as $\mathsf{g}^\alpha_{\ \mu}\mathsf{g}^\beta_{\ \nu}\mathsf{g}_{\alpha\beta}$.

On a formal level the metric $\mathsf{g}_{\alpha\beta}$ is obtained by solving the Einstein field equations, and the world line is determined by the equations of energy-momentum conservation, which follow from the field equations. From Eqs. (5.14), (13.3), and (19.3) we obtain

$$\nabla_\beta \mathsf{T}^{\alpha\beta} = m \int_\gamma \frac{d}{d\lambda}\biggl(\frac{\mathsf{g}^\alpha_{\ \mu}\dot{z}^\mu}{\sqrt{-\mathsf{g}_{\mu\nu}\dot{z}^\mu \dot{z}^\nu}} \biggr) \delta_4(x,z)\, d\lambda,$$

and additional manipulations reduce this to

$$\nabla_\beta \mathsf{T}^{\alpha\beta} = m \int_\gamma \frac{\mathsf{g}^\alpha_{\ \mu}}{\sqrt{-\mathsf{g}_{\mu\nu}\dot{z}^\mu \dot{z}^\nu}} \biggl(\frac{\mathsf{D}\dot{z}^\mu}{d\lambda} - \mathsf{k}\dot{z}^\mu \biggr) \delta_4(x,z)\, d\lambda, \tag{19.4}$$

where $\mathsf{D}\dot{z}^\mu/d\lambda$ is the covariant acceleration and k is a scalar field on the world line. Energy-momentum conservation therefore produces the geodesic equation

$$\frac{\mathsf{D}\dot{z}^\mu}{d\lambda} = \mathsf{k}\dot{z}^\mu, \tag{19.5}$$

and

$$\mathsf{k} := \frac{1}{\sqrt{-\mathsf{g}_{\mu\nu}\dot{z}^\mu\dot{z}^\nu}}\frac{d}{d\lambda}\sqrt{-\mathsf{g}_{\mu\nu}\dot{z}^\mu\dot{z}^\nu} \tag{19.6}$$

measures the failure of λ to be an affine parameter on the geodesic γ.

Decomposition into background and perturbation

At this stage we begin treating m as a small quantity, and we write

$$\mathsf{g}_{\alpha\beta} = g_{\alpha\beta} + h_{\alpha\beta}, \tag{19.7}$$

with $g_{\alpha\beta}$ denoting the $m \to 0$ limit of the metric $\mathsf{g}_{\alpha\beta}$, and $h_{\alpha\beta}$ containing the dependence on m. We shall refer to $g_{\alpha\beta}$ as the "metric of the background spacetime" and to $h_{\alpha\beta}$ as the "perturbation" produced by the particle. We insist, however, that no approximation is introduced at this stage; the perturbation $h_{\alpha\beta}$ is the *exact difference* between the exact metric $\mathsf{g}_{\alpha\beta}$ and the background metric $g_{\alpha\beta}$. Below we shall use the background metric to lower and raise indices.

We introduce the tensor field

$$C^\alpha_{\beta\gamma} := \mathsf{\Gamma}^\alpha_{\beta\gamma} - \Gamma^\alpha_{\beta\gamma} \tag{19.8}$$

as the *exact difference* between $\mathsf{\Gamma}^\alpha_{\beta\gamma}$, the connection compatible with the exact metric $\mathsf{g}_{\alpha\beta}$, and $\Gamma^\alpha_{\beta\gamma}$, the connection compatible with the background metric $g_{\alpha\beta}$. A covariant differentiation indicated by $;\alpha$ will refer to $\Gamma^\alpha_{\beta\gamma}$, while a covariant differentiation indicated by ∇_α will continue to refer to $\mathsf{\Gamma}^\alpha_{\beta\gamma}$.

We express the exact Einstein tensor as

$$\mathsf{G}^{\alpha\beta} = G^{\alpha\beta}[g] + \delta G^{\alpha\beta}[g,h] + \Delta G^{\alpha\beta}[g,h], \tag{19.9}$$

where $G^{\alpha\beta}$ is the Einstein tensor of the background spacetime, which is assumed to vanish. The second term $\delta G^{\alpha\beta}$ is the *linearized Einstein operator* defined by

$$\delta G^{\alpha\beta} := -\frac{1}{2}\big(\Box\gamma^{\alpha\beta} + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}\gamma^{\gamma\delta}\big) + \frac{1}{2}\big(\gamma^{\alpha\gamma}{}_{;\gamma}{}^{\beta} + \gamma^{\beta\gamma}{}_{;\gamma}{}^{\alpha} - g^{\alpha\beta}\gamma^{\gamma\delta}{}_{;\gamma\delta}\big), \tag{19.10}$$

where $\Box\gamma^{\alpha\beta} := g^{\gamma\delta}\gamma^{\alpha\beta}{}_{;\gamma\delta}$ is the wave operator in the background spacetime, and

$$\gamma^{\alpha\beta} := h^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}\big(g_{\gamma\delta}h^{\gamma\delta}\big) \tag{19.11}$$

is the "trace-reversed" metric perturbation (with all indices raised with the background metric). The third term $\Delta G^{\alpha\beta}$ contains the remaining nonlinear pieces that are excluded from $\delta G^{\alpha\beta}$.

Field equations and conservation statement

The exact Einstein field equations can be expressed as

$$\delta G^{\alpha\beta} = 8\pi T^{\alpha\beta}_{\text{eff}}, \tag{19.12}$$

where the effective energy-momentum tensor is defined by

$$T^{\alpha\beta}_{\text{eff}} := \mathsf{T}^{\alpha\beta} - \frac{1}{8\pi}\Delta G^{\alpha\beta}. \tag{19.13}$$

Because $\delta G^{\alpha\beta}$ satisfies the Bianchi-like identities $\delta G^{\alpha\beta}{}_{;\beta} = 0$, the effective energy-momentum tensor is conserved in the background spacetime:

$$T^{\alpha\beta}_{\text{eff}\,;\beta} = 0. \tag{19.14}$$

This statement is equivalent to $\nabla_\beta \mathsf{T}^{\alpha\beta} = 0$, as can be inferred from the equations $\nabla_\beta\mathsf{G}^{\alpha\beta} = \mathsf{G}^{\alpha\beta}{}_{;\beta} + C^\alpha_{\gamma\beta}\mathsf{G}^{\gamma\beta} + C^\beta_{\gamma\beta}\mathsf{G}^{\alpha\gamma}$, $\nabla_\beta\mathsf{T}^{\alpha\beta} = \mathsf{T}^{\alpha\beta}{}_{;\beta} + C^\alpha_{\gamma\beta}\mathsf{T}^{\gamma\beta} + C^\beta_{\gamma\beta}\mathsf{T}^{\alpha\gamma}$, and the definition of $T^{\alpha\beta}_{\text{eff}}$. Equation (19.14), in turn, is equivalent to Eq. (19.5), which states that the motion of the point particle is geodesic in the perturbed spacetime.

Integration of the field equations

Equation (19.12) expresses the full and exact content of Einstein's field equations. It is written in such a way that the left-hand side is linear in the perturbation $h_{\alpha\beta}$, while the right-hand side contains all nonlinear terms. It may be viewed formally as a set of linear differential equations for $h_{\alpha\beta}$ with a specified source term $T^{\alpha\beta}_{\rm eff}$. This equation is of mixed hyperbolic-elliptic type, and as such it is a poor starting point for the selection of retarded solutions that enforce a strict causal link between the source and the field. This inadequacy, however, can be remedied by imposing the *Lorenz gauge condition*

$$\gamma^{\alpha\beta}{}_{;\beta} = 0, \tag{19.15}$$

which converts $\delta G^{\alpha\beta}$ into a strictly hyperbolic differential operator. In this gauge the field equations become

$$\Box\gamma^{\alpha\beta} + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}\gamma^{\gamma\delta} = -16\pi T^{\alpha\beta}_{\rm eff}. \tag{19.16}$$

This is a tensorial wave equation formulated in the background spacetime, and while the left-hand side is manifestly linear in $h_{\alpha\beta}$, the right-hand side continues to incorporate all nonlinear terms. Equations (19.15) and (19.16) still express the full content of the exact field equations.

A formal solution to Eq. (19.16) is

$$\gamma^{\alpha\beta}(x) = 4\int G_{+\ \gamma'\delta'}^{\ \alpha\beta}(x,x')T^{\gamma'\delta'}_{\rm eff}(x')\sqrt{-g'}\, d^4x', \tag{19.17}$$

where $G_{+\ \gamma'\delta'}^{\ \alpha\beta}(x,x')$ is the retarded Green's function introduced in Sec. 16. With the help of Eq. (16.21), it is easy to show that

$$\gamma^{\alpha\beta}{}_{;\beta} = 4\int G_{+\gamma'}^{\ \alpha} T^{\gamma'\delta'}_{\rm eff\ ;\delta'}\sqrt{-g'}\, d^4x' \tag{19.18}$$

follows directly from Eq. (19.17); $G_{+\ \gamma'}^{\ \alpha}(x,x')$ is the electromagnetic Green's function introduced in Sec. 15. This equation indicates that the Lorenz gauge condition is automatically enforced when the conservation equation $T^{\alpha\beta}_{\rm eff\,;\beta} = 0$ is imposed. Conversely, Eq. (19.18) implies that $\Box(\gamma^{\alpha\beta}{}_{;\beta}) = -16\pi T^{\alpha\beta}_{\rm eff\,;\beta}$, which indicates that imposition of $\gamma^{\alpha\beta}{}_{;\beta} = 0$ automatically enforces the conservation equation. There is a one-to-one correspondence between the conservation equation and the Lorenz gauge condition.

The split of the Einstein field equations into a wave equation and a gauge condition directly tied to the conservation of the effective energy-momentum tensor is a most powerful tool, because it allows us to disentangle the problems of obtaining $h_{\alpha\beta}$ and determining the motion of the particle. This comes about because the wave equation can be solved first, independently of the gauge condition, for a particle moving on an arbitrary world line γ; the world line is determined next, by imposing the Lorenz gauge condition on the solution to the wave equation. More precisely stated, the source term $T^{\alpha\beta}_{\rm eff}$ for the wave equation can be evaluated for any world line γ, without demanding that the effective energy-momentum tensor be conserved, and without demanding that γ be a geodesic of the perturbed spacetime. Solving the wave equation then returns $h_{\alpha\beta}[\gamma]$ as a functional of the arbitrary world line, and the metric is not yet fully specified. Because imposing the Lorenz gauge condition is equivalent to imposing conservation of the effective energy-momentum tensor, inserting $h_{\alpha\beta}[\gamma]$ within Eq. (19.15) finally determines γ, and forces it to be a geodesic of the perturbed spacetime. At this stage the full set of Einstein field equations is accounted for, and the metric is fully specified as a tensor field in spacetime. The split of the field equations into a wave equation and a gauge condition is key to the formulation of the gravitational self-force; in this specific context the Lorenz gauge is conferred a preferred status among all choices of gauge.

An important question to be addressed is how the wave equation is to be integrated. A method of principle, based on the assumed smallness of m and $h_{\alpha\beta}$, is suggested by post-Minkowskian theory [143, 144]. One proceeds by iterations. In the first iterative stage, one fixes γ and substitutes $h_0^{\alpha\beta} = 0$ within $T^{\alpha\beta}_{\rm eff}$; evaluation of the integral in Eq. (19.17) returns the first-order approximation $h_1^{\alpha\beta}[\gamma] = O(m)$ for the perturbation. In the second stage $h_1^{\alpha\beta}$ is inserted within $T^{\alpha\beta}_{\rm eff}$ and Eq. (19.17) returns the second-order approximation $h_2^{\alpha\beta}[\gamma] = O(m, m^2)$ for the perturbation. Assuming that this procedure can be repeated at will and produces an adequate asymptotic series for the exact perturbation, the iterations are stopped when the $n^{\rm th}$-order approximation $h_n^{\alpha\beta}[\gamma] = O(m, m^2, \cdots, m^n)$ is deemed to be sufficiently accurate. The world

line is then determined, to order m^n, by subjecting the approximated field to the Lorenz gauge condition. It is to be noted that the procedure necessarily produces an approximation of the field, and an approximation of the motion, because the number of iterations is necessarily finite. This is the *only source of approximation* in our formulation of the dynamics of a point mass.

Equations of motion

Conservation of energy-momentum implies Eq. (19.5), which states that the motion of the point mass is geodesic in the perturbed spacetime. The equation is expressed in terms of the exact connection $\mathsf{\Gamma}^{\alpha}_{\beta\gamma}$, and with the help of Eq. (19.8) it can be re-written in terms of the background connection $\Gamma^{\alpha}_{\beta\gamma}$. We get $D\dot{z}^\mu/d\lambda = -C^{\mu}_{\nu\lambda}\dot{z}^\nu\dot{z}^\lambda + \mathsf{k}\dot{z}^\mu$, where the left-hand side is the covariant acceleration in the background spacetime, and k is given by Eq. (19.6). At this stage the arbitrary parameter λ can be identified with proper time τ in the background spacetime. With this choice the equations of motion become

$$a^\mu = -C^{\mu}_{\nu\lambda}u^\nu u^\lambda + \mathsf{k}u^\mu, \tag{19.19}$$

where $u^\mu := dz^\mu/d\tau$ is the velocity vector in the background spacetime, $a^\mu := Du^\mu/d\tau$ the covariant acceleration, and

$$\mathsf{k} = \frac{1}{\sqrt{1-h_{\mu\nu}u^\mu u^\nu}}\frac{d}{d\tau}\sqrt{1-h_{\mu\nu}u^\mu u^\nu}. \tag{19.20}$$

Equation (19.19) is an exact statement of the equations of motion. It expresses the fact that while the motion is geodesic in the perturbed spacetime, it may be viewed as accelerated motion in the background spacetime. Because $h_{\alpha\beta}$ is calculated as an expansion in powers of m, the acceleration also is eventually obtained as an expansion in powers of m. Here, in keeping with the preceding sections, we will use order-reduction to make that expansion well-behaved; in Part V of the review, we will formulate the expansion more clearly as part of more systematic approach.

Implementation to first order in m

While our formulation of the dynamics of a point mass is in principle exact, any practical implementation will rely on an approximation method. As we saw previously, the most immediate source of approximation concerns the number of iterations involved in the integration of the wave equation. Here we perform a single iteration and obtain the perturbation $h_{\alpha\beta}$ and the equations of motion to first order in the mass m.

In a first iteration of the wave equation we fix γ and set $\Delta G^{\alpha\beta} = 0$, $\mathsf{T}^{\alpha\beta} = T^{\alpha\beta}$, where

$$T^{\alpha\beta} = m\int_\gamma g^{\alpha}_{\ \mu}(x,z)g^{\beta}_{\ \nu}(x,z)u^\mu u^\nu\,\delta_4(x,z)\,d\tau \tag{19.21}$$

is the particle's energy-momentum tensor in the background spacetime. This implies that $T^{\alpha\beta}_{\mathrm{eff}} = T^{\alpha\beta}$, and Eq. (19.16) becomes

$$\Box\gamma^{\alpha\beta} + 2R_{\gamma\ \delta}^{\ \alpha\ \beta}\gamma^{\gamma\delta} = -16\pi T^{\alpha\beta} + O(m^2). \tag{19.22}$$

Its solution is

$$\gamma^{\alpha\beta}(x) = 4m\int_\gamma G_{+\ \mu\nu}^{\ \alpha\beta}(x,z)u^\mu u^\nu\,d\tau + O(m^2), \tag{19.23}$$

and from this we obtain $h^{\alpha\beta}$. Equation (19.8) gives rise to $C^{\alpha}_{\beta\gamma} = \frac{1}{2}(h^{\alpha}_{\ \beta;\gamma} + h^{\alpha}_{\ \gamma;\beta} - h_{\beta\gamma}^{\ \ \ ;\alpha}) + O(m^2)$, and from Eq. (19.20) we obtain $\mathsf{k} = -\frac{1}{2}h_{\nu\lambda;\rho}u^\nu u^\lambda u^\rho - h_{\nu\lambda}u^\nu a^\lambda + O(m^2)$; we can discard the second term because it is clear that the acceleration will be of order m. Inserting these results within Eq. (19.19), we obtain

$$a^\mu = -\frac{1}{2}\Big(h^{\mu}_{\ \nu;\lambda} + h^{\mu}_{\ \lambda;\nu} - h_{\nu\lambda}^{\ \ \ ;\mu} + u^\mu h_{\nu\lambda;\rho}u^\rho\Big)u^\nu u^\lambda + O(m^2). \tag{19.24}$$

We express this in the equivalent form

$$a^\mu = -\frac{1}{2}\big(g^{\mu\nu} + u^\mu u^\nu\big)\big(2h_{\nu\lambda;\rho} - h_{\lambda\rho;\nu}\big)u^\lambda u^\rho + O(m^2) \tag{19.25}$$

to emphasize the fact that the acceleration is orthogonal to the velocity vector.

It should be clear that Eq. (19.25) is valid only in a formal sense, because the potentials obtained from Eqs. (19.23) diverge on the world line. To make sense of these equations we will proceed as in Secs. 17 and 18 with a careful analysis of the field's singularity structure; regularization will produce a well-defined version of Eq. (19.25). Our formulation of the dynamics of a point mass makes it clear that a proper implementation requires that the wave equation of Eq. (19.22) and the equations of motion of Eq. (19.25) be integrated simultaneously, in a self-consistent manner.

Failure of a strictly linearized formulation

In the preceding discussion we started off with an exact formulation of the problem of motion for a small mass m in a background spacetime with metric $g_{\alpha\beta}$, but eventually boiled it down to an implementation accurate to first order in m. Would it not be simpler and more expedient to formulate the problem directly to first order? The answer is a resounding no: By doing so we would be driven toward a grave inconsistency; the nonlinear formulation is absolutely necessary if one wishes to contemplate a self-consistent integration of Eqs. (19.22) and (19.25).

A strictly linearized formulation of the problem would be based on the field equations $\delta G^{\alpha\beta} = 8\pi T^{\alpha\beta}$, where $T^{\alpha\beta}$ is the energy-momentum tensor of Eq. (19.21). The Bianchi-like identities $\delta G^{\alpha\beta}{}_{;\beta} = 0$ dictate that $T^{\alpha\beta}$ must be *conserved in the background spacetime*, and a calculation identical to the one leading to Eq. (19.5) would reveal that the particle's motion must be *geodesic in the background spacetime*. In the strictly linearized formulation, therefore, the gravitational potentials of Eq. (19.23) must be sourced by a particle moving on a geodesic, and there is no opportunity for these potentials to exert a self-force. To get the self-force, one must provide a formulation that extends beyond linear order. To be sure, one could persist in adopting the linearized formulation and "save the phenomenon" by relaxing the conservation equation. In practice this could be done by adopting the solutions of Eq. (19.23), demanding that the motion be geodesic in the perturbed spacetime, and relaxing the linearized gauge condition to $\gamma^{\alpha\beta}{}_{;\beta} = O(m^2)$. While this prescription would produce the correct answer, it is largely *ad hoc* and does not come with a clear justification. Our exact formulation provides much more control, at least in a formal sense. We shall do even better in Part V.

An alternative formulation of the problem provided by Gralla and Wald [16] avoids the inconsistency by refraining from performing a self-consistent integration of Eqs. (19.22) and (19.25). Instead of an expansion of the acceleration in powers of m, their approach is based on an expansion of the world line itself, and it returns the equations of motion for a deviation vector which describes the offset of the true world line relative to a reference geodesic. While this approach is mathematically sound, it eventually breaks down as the deviation vector becomes large, and it does not provide a justification of the self-consistent treatment of the equations.

The difference between the Gralla-Wald approach and a self-consistent one is the difference between a regular expansion and a general one. In a regular expansion, all dependence on a small quantity m is expanded in powers:

$$h_{\alpha\beta}(x,m) = \sum_{n=0}^{\infty} m^n h^{(n)}_{\alpha\beta}(x). \tag{19.26}$$

In a general expansion, on the other hand, the functions $h^{(n)}_{\alpha\beta}$ are allowed to retain some dependence on the small quantity:

$$h_{\alpha\beta}(x,m) = \sum_{n=0}^{\infty} m^n h^{(n)}_{\alpha\beta}(x,m). \tag{19.27}$$

Put simply, the goal of a general expansion is to expand only *part* of a function's dependence on a small quantity, while holding fixed some specific dependence that captures one or more of the function's essential features. In the self-consistent expansion that we advocate here, our iterative solution returns

$$h^{N}_{\alpha\beta}(x,m) = \sum_{n=0}^{N} m^n h^{(n)}_{\alpha\beta}(x;\gamma(m)), \tag{19.28}$$

in which the functional dependence on the world line γ incorporates a dependence on the expansion parameter m. We deliberately introduce this functional dependence on a mass-dependent world line in order to maintain a meaningful and accurate description of the particle's motion. Although the regular expansion can be retrieved by further expanding the dependence within $\gamma(m)$, the reverse statement does not hold: the general expansion cannot be justified on the basis of the regular one. The notion of a general expansion is at the core of singular perturbation theory [18, 131, 145–148]. We shall return to these issues in our treatment of asymptotically small bodies, and in particular, in Sec. 22.5 below.

Vacuum background spacetime

To conclude this subsection we should explain why it is desirable to restrict our discussion to spacetimes that contain no matter except for the point particle. Suppose, in contradiction with this assumption, that the background spacetime contains a distribution of matter around which the particle is moving. (The corresponding vacuum situation has the particle moving around a black hole. Notice that we are still assuming that the particle moves in a region of spacetime in which there is no matter; the issue is whether we can allow for a distribution of matter *somewhere else.*) Suppose also that the matter distribution is described by a collection of matter fields Ψ. Then the field equations satisfied by the matter have the schematic form $E[\Psi; g] = 0$, and the background metric is determined by the Einstein field equations $G[g] = 8\pi M[\Psi; g]$, in which $M[\Psi; g]$ stands for the matter's energy-momentum tensor. We now insert the point particle in the spacetime, and recognize that this displaces the background solution (Ψ, g) to a new solution $(\Psi + \delta\Psi, g + \delta g)$. The perturbations are determined by the coupled set of equations $E[\Psi + \delta\Psi; g + \delta g] = 0$ and $G[g + \delta g] = 8\pi M[\Psi + \delta\Psi; g + \delta g] + 8\pi T[g]$. After linearization these take the form of

$$E_\Psi \cdot \delta\Psi + E_g \cdot \delta g = 0, \qquad G_g \cdot \delta g = 8\pi\bigl(M_\Psi \cdot \delta\Psi + M_g \cdot \delta g + T\bigr),$$

where E_Ψ, E_g, M_Φ, and M_g are suitable differential operators acting on the perturbations. This is a *coupled set* of partial differential equations for the perturbations $\delta\Psi$ and δg. These equations are linear, but they are much more difficult to deal with than the single equation for δg that was obtained in the vacuum case. And although it is still possible to solve the coupled set of equations via a Green's function technique, the degree of difficulty is such that we will not attempt this here. We shall, therefore, continue to restrict our attention to the case of a point particle moving in a vacuum (globally Ricci-flat) background spacetime.

19.2 Retarded potentials near the world line

Going back to Eq. (19.23), we have that the gravitational potentials associated with a point particle of mass m moving on world line γ are given by

$$\gamma^{\alpha\beta}(x) = 4m \int_\gamma G_{+\ \mu\nu}^{\ \alpha\beta}(x, z) u^\mu u^\nu \, d\tau, \tag{19.29}$$

up to corrections of order m^2; here $z^\mu(\tau)$ gives the description of the world line, $u^\mu = dz^\mu/d\tau$ is the velocity vector, and $G_{+\ \gamma'\delta'}^{\ \alpha\beta}(x, x')$ is the retarded Green's function introduced in Sec. 16. Because the retarded Green's function is defined globally in the entire background spacetime, Eq. (19.29) describes the gravitational perturbation created by the particle at any point x in that spacetime.

For a more concrete expression we take x to be in a neighbourhood of the world line. The manipulations that follow are very close to those performed in Sec. 17.2 for the case of a scalar charge, and in Sec. 18.2 for the case of an electric charge. Because these manipulations are by now familiar, it will be sufficient here to present only the main steps. There are two important simplifications that occur in the case of a massive particle. First, it is clear that

$$a^\mu = O(m) = \dot{a}^\mu, \tag{19.30}$$

and we will take the liberty of performing a pre-emptive order-reduction by dropping all terms involving the acceleration vector when computing $\gamma^{\alpha\beta}$ and $\gamma_{\alpha\beta;\gamma}$ to first order in m; otherwise we would arrive at an equation for the acceleration that would include an antidamping term $-\frac{11}{3}m\dot{a}^\mu$ [7, 149, 150]. Second, because we take $g_{\alpha\beta}$ to be a solution to the vacuum field equations, we are also allowed to set

$$R_{\mu\nu}(z) = 0 \tag{19.31}$$

in our computations.

With the understanding that x is close to the world line (refer back to Fig. 9), we substitute the Hadamard construction of Eq. (16.7) into Eq. (19.29) and integrate over the portion of γ that is contained in $\mathcal{N}(x)$. The result is

$$\gamma^{\alpha\beta}(x) = \frac{4m}{r} U^{\alpha\beta}{}_{\gamma'\delta'}(x,x') u^{\gamma'} u^{\delta'} + 4m \int_{\tau_<}^{u} V^{\alpha\beta}{}_{\mu\nu}(x,z) u^\mu u^\nu \, d\tau + 4m \int_{-\infty}^{\tau_<} G_{+}{}^{\alpha\beta}{}_{\mu\nu}(x,z) u^\mu u^\nu \, d\tau, \tag{19.32}$$

in which primed indices refer to the retarded point $x' := z(u)$ associated with x, $r := \sigma_{\alpha'} u^{\alpha'}$ is the retarded distance from x' to x, and $\tau_<$ is the proper time at which γ enters $\mathcal{N}(x)$ from the past.

In the following subsections we shall refer to $\gamma_{\alpha\beta}(x)$ as the *gravitational potentials* at x produced by a particle of mass m moving on the world line γ, and to $\gamma_{\alpha\beta;\gamma}(x)$ as the *gravitational field* at x. To compute this is our next task.

19.3 Gravitational field in retarded coordinates

Keeping in mind that x' and x are related by $\sigma(x,x') = 0$, a straightforward computation reveals that the covariant derivatives of the gravitational potentials are given by

$$\begin{aligned}
\gamma_{\alpha\beta;\gamma}(x) &= -\frac{4m}{r^2} U_{\alpha\beta\alpha'\beta'} u^{\alpha'} u^{\beta'} \partial_\gamma r + \frac{4m}{r} U_{\alpha\beta\alpha'\beta';\gamma} u^{\alpha'} u^{\beta'} + \frac{4m}{r} U_{\alpha\beta\alpha'\beta';\gamma'} u^{\alpha'} u^{\beta'} u^{\gamma'} \partial_\gamma u \\
&\quad + 4m V_{\alpha\beta\alpha'\beta'} u^{\alpha'} u^{\beta'} \partial_\gamma u + \gamma^{\rm tail}_{\alpha\beta\gamma}(x),
\end{aligned} \tag{19.33}$$

where the "tail integral" is defined by

$$\begin{aligned}
\gamma^{\rm tail}_{\alpha\beta\gamma}(x) &= 4m \int_{\tau_<}^{u} \nabla_\gamma V_{\alpha\beta\mu\nu}(x,z) u^\mu u^\nu \, d\tau + 4m \int_{-\infty}^{\tau_<} \nabla_\gamma G_{+\alpha\beta\mu\nu}(x,z) u^\mu u^\nu \, d\tau \\
&= 4m \int_{-\infty}^{u^-} \nabla_\gamma G_{+\alpha\beta\mu\nu}(x,z) u^\mu u^\nu \, d\tau.
\end{aligned} \tag{19.34}$$

The second form of the definition, in which the integration is cut short at $\tau = u^- := u - 0^+$ to avoid the singular behaviour of the retarded Green's function at $\sigma = 0$, is equivalent to the first form.

We wish to express $\gamma_{\alpha\beta;\gamma}(x)$ in the retarded coordinates of Sec. 10, as an expansion in powers of r. For this purpose we decompose the field in the tetrad (e^α_0, e^α_a) that is obtained by parallel transport of $(u^{\alpha'}, e^{\alpha'}_a)$ on the null geodesic that links x to x'; this construction is detailed in Sec. 10. We recall from Eq. (10.4) that the parallel propagator can be expressed as $g^{\alpha'}{}_{\alpha} = u^{\alpha'} e^0_\alpha + e^{\alpha'}_a e^a_\alpha$. The expansion relies on Eq. (10.29) for $\partial_\gamma u$ and Eq. (10.31) for $\partial_\gamma r$, both simplified by setting $a_a = 0$. We shall also need

$$U_{\alpha\beta\alpha'\beta'} u^{\alpha'} u^{\beta'} = g^{\alpha'}{}_{(\alpha} g^{\beta'}{}_{\beta)} \Big[u_{\alpha'} u_{\beta'} + O(r^3) \Big], \tag{19.35}$$

which follows from Eq. (16.13),

$$\begin{aligned}
U_{\alpha\beta\alpha'\beta';\gamma} u^{\alpha'} u^{\beta'} &= g^{\alpha'}{}_{(\alpha} g^{\beta'}{}_{\beta)} g^{\gamma'}{}_{\gamma} \Big[-r \big(R_{\alpha'0\gamma'0} + R_{\alpha'0\gamma'd} \Omega^d \big) u_{\beta'} + O(r^2) \Big], && (19.36) \\
U_{\alpha\beta\alpha'\beta';\gamma'} u^{\alpha'} u^{\beta'} u^{\gamma'} &= g^{\alpha'}{}_{(\alpha} g^{\beta'}{}_{\beta)} \Big[r R_{\alpha'0d0} \Omega^d u_{\beta'} + O(r^2) \Big], && (19.37)
\end{aligned}$$

which follow from Eqs. (16.14) and (16.15), respectively, as well as the relation $\sigma^{\alpha'} = -r(u^{\alpha'} + \Omega^a e^{\alpha'}_a)$ first encountered in Eq. (10.7). And finally, we shall need

$$V_{\alpha\beta\alpha'\beta'} u^{\alpha'} u^{\beta'} = g^{\alpha'}{}_{(\alpha} g^{\beta'}{}_{\beta)} \Big[R_{\alpha'0\beta'0} + O(r) \Big], \tag{19.38}$$

which follows from Eq. (16.17).

Making these substitutions in Eq. (19.3) and projecting against various members of the tetrad gives

$$\begin{aligned}
\gamma_{000}(u,r,\Omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)e^\alpha_0(x)e^\beta_0(x)e^\gamma_0(x) = 2mR_{a0b0}\Omega^a\Omega^b + \gamma^{\rm tail}_{000} + O(r), && (19.39)\\
\gamma_{0b0}(u,r,\Omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)e^\alpha_0(x)e^\beta_b(x)e^\gamma_0(x) = -4mR_{b0c0}\Omega^c + \gamma^{\rm tail}_{0b0} + O(r), && (19.40)\\
\gamma_{ab0}(u,r,\Omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)e^\alpha_a(x)e^\beta_b(x)e^\gamma_0(x) = 4mR_{a0b0} + \gamma^{\rm tail}_{ab0} + O(r), && (19.41)\\
\gamma_{00c}(u,r,\Omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)e^\alpha_0(x)e^\beta_0(x)e^\gamma_c(x)\\
&= -4m\biggl[\Bigl(\frac{1}{r^2} + \frac{1}{3}R_{a0b0}\Omega^a\Omega^b\Bigr)\Omega_c + \frac{1}{6}R_{c0b0}\Omega^b - \frac{1}{6}R_{ca0b}\Omega^a\Omega^b\biggr] + \gamma^{\rm tail}_{00c} + O(r), && (19.42)\\
\gamma_{0bc}(u,r,\Omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)e^\alpha_0(x)e^\beta_b(x)e^\gamma_c(x)\\
&= 2m\bigl(R_{b0c0} + R_{b0cd}\Omega^d + R_{b0d0}\Omega^d\Omega_c\bigr) + \gamma^{\rm tail}_{0bc} + O(r), && (19.43)\\
\gamma_{abc}(u,r,\Omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)e^\alpha_a(x)e^\beta_b(x)e^\gamma_c(x) = -4mR_{a0b0}\Omega_c + \gamma^{\rm tail}_{abc} + O(r), && (19.44)
\end{aligned}$$

where, for example, $R_{a0b0}(u) := R_{\alpha'\gamma'\beta'\delta'}e^{\alpha'}_a u^{\gamma'} e^{\beta'}_b u^{\delta'}$ are frame components of the Riemann tensor evaluated at $x' := z(u)$. We have also introduced the frame components of the tail part of the gravitational field, which are obtained from Eq. (19.34) evaluated at x' instead of x; for example, $\gamma^{\rm tail}_{000} = u^{\alpha'}u^{\beta'}u^{\gamma'}\gamma^{\rm tail}_{\alpha'\beta'\gamma'}(x')$. We may note here that while γ_{00c} is the only component of the gravitational field that diverges when $r \to 0$, the other components are nevertheless singular because of their dependence on the unit vector Ω^a; the only exception is γ_{ab0}, which is regular.

19.4 Gravitational field in Fermi normal coordinates

The translation of the results contained in Eqs. (19.39)–(19.44) into the Fermi normal coordinates of Sec. 9 proceeds as in Secs. 17.4 and 18.4, but is simplified by setting $a_a = \dot{a}_0 = \dot{a}_a = 0$ in Eqs. (11.7), (11.8), (11.4), (11.5), and (11.6) that relate the Fermi normal coordinates (t,s,ω^a) to the retarded coordinates. We recall that the Fermi normal coordinates refer to a point $\bar{x} := z(t)$ on the world line that is linked to x by a spacelike geodesic that intersects γ orthogonally.

The translated results are

$$\begin{aligned}
\bar\gamma_{000}(t,s,\omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)\bar{e}^\alpha_0(x)\bar{e}^\beta_0(x)\bar{e}^\gamma_0(x) = \bar\gamma^{\rm tail}_{000} + O(s), && (19.45)\\
\bar\gamma_{0b0}(t,s,\omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)\bar{e}^\alpha_0(x)\bar{e}^\beta_b(x)\bar{e}^\gamma_0(x) = -4mR_{b0c0}\omega^c + \bar\gamma^{\rm tail}_{0b0} + O(s), && (19.46)\\
\bar\gamma_{ab0}(t,s,\omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)\bar{e}^\alpha_a(x)\bar{e}^\beta_b(x)\bar{e}^\gamma_0(x) = 4mR_{a0b0} + \bar\gamma^{\rm tail}_{ab0} + O(s), && (19.47)\\
\bar\gamma_{00c}(t,s,\omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)\bar{e}^\alpha_0(x)\bar{e}^\beta_0(x)\bar{e}^\gamma_c(x)\\
&= -4m\biggl[\Bigl(\frac{1}{s^2} - \frac{1}{6}R_{a0b0}\omega^a\omega^b\Bigr)\omega_c + \frac{1}{3}R_{c0b0}\omega^b\biggr] + \bar\gamma^{\rm tail}_{00c} + O(s), && (19.48)\\
\bar\gamma_{0bc}(t,s,\omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)\bar{e}^\alpha_0(x)\bar{e}^\beta_b(x)\bar{e}^\gamma_c(x) = 2m\bigl(R_{b0c0} + R_{b0cd}\omega^d\bigr) + \bar\gamma^{\rm tail}_{0bc} + O(s), && (19.49)\\
\bar\gamma_{abc}(t,s,\omega^a) &:= \gamma_{\alpha\beta;\gamma}(x)\bar{e}^\alpha_a(x)\bar{e}^\beta_b(x)\bar{e}^\gamma_c(x) = -4mR_{a0b0}\omega_c + \bar\gamma^{\rm tail}_{abc} + O(s), && (19.50)
\end{aligned}$$

where all frame components are now evaluated at $\bar{x}$ instead of x'.

It is then a simple matter to average these results over a two-surface of constant t and s. Using the area element of Eq. (17.24) and definitions analogous to those of Eq. (17.25), we obtain

$$\begin{aligned}
\langle\bar\gamma_{000}\rangle &= \bar\gamma^{\rm tail}_{000} + O(s), && (19.51)\\
\langle\bar\gamma_{0b0}\rangle &= \bar\gamma^{\rm tail}_{0b0} + O(s), && (19.52)\\
\langle\bar\gamma_{ab0}\rangle &= 4mR_{a0b0} + \bar\gamma^{\rm tail}_{ab0} + O(s), && (19.53)\\
\langle\bar\gamma_{00c}\rangle &= \bar\gamma^{\rm tail}_{00c} + O(s), && (19.54)\\
\langle\bar\gamma_{0bc}\rangle &= 2mR_{b0c0} + \bar\gamma^{\rm tail}_{0bc} + O(s), && (19.55)\\
\langle\bar\gamma_{abc}\rangle &= \bar\gamma^{\rm tail}_{abc} + O(s). && (19.56)
\end{aligned}$$

The averaged gravitational field is regular in the limit $s \to 0$, in which the tetrad $(\bar{e}^\alpha_0, \bar{e}^\alpha_a)$ coincides with $(u^{\bar{\alpha}}, e^{\bar{\alpha}}_a)$. Reconstructing the field at $\bar{x}$ from its frame components gives

$$\langle \gamma_{\bar{\alpha}\bar{\beta};\bar{\gamma}} \rangle = -4m\Bigl(u_{(\bar{\alpha}} R_{\bar{\beta})\bar{\delta}\bar{\gamma}\bar{\epsilon}} + R_{\bar{\alpha}\bar{\delta}\bar{\beta}\bar{\epsilon}} u_{\bar{\gamma}} \Bigr) u^{\bar{\delta}} u^{\bar{\epsilon}} + \gamma^{\rm tail}_{\bar{\alpha}\bar{\beta}\bar{\gamma}}, \tag{19.57}$$

where the tail term can be copied from Eq. (19.34),

$$\gamma^{\rm tail}_{\bar{\alpha}\bar{\beta}\bar{\gamma}}(\bar{x}) = 4m \int_{-\infty}^{t^-} \nabla_{\bar{\gamma}} G_{+\bar{\alpha}\bar{\beta}\mu\nu}(\bar{x}, z) u^\mu u^\nu \, d\tau. \tag{19.58}$$

The tensors that appear in Eq. (19.57) all refer to the simultaneous point $\bar{x} := z(t)$, which can now be treated as an arbitrary point on the world line γ.

19.5 Singular and regular fields

The singular gravitational potentials

$$\gamma^{\alpha\beta}_{\rm S}(x) = 4m \int_\gamma G_{{\rm S}\ \ \mu\nu}^{\ \alpha\beta}(x, z) u^\mu u^\nu \, d\tau \tag{19.59}$$

are solutions to the wave equation of Eq. (19.22); the singular Green's function was introduced in Sec. 16.5. We will see that the singular field $\gamma^{\rm S}_{\alpha\beta;\gamma}$ reproduces the singular behaviour of the retarded solution near the world line, and that the difference, $\gamma^{\rm R}_{\alpha\beta;\gamma} = \gamma_{\alpha\beta;\gamma} - \gamma^{\rm S}_{\alpha\beta;\gamma}$, is regular on the world line.

To evaluate the integral of Eq. (19.59) we take x to be close to the world line (see Fig. 9), and we invoke Eq. (16.31) as well as the Hadamard construction of Eq. (16.37). This gives

$$\gamma^{\alpha\beta}_{\rm S}(x) = \frac{2m}{r} U^{\alpha\beta}_{\ \ \ \gamma'\delta'} u^{\gamma'} u^{\delta'} + \frac{2m}{r_{\rm adv}} U^{\alpha\beta}_{\ \ \ \gamma''\delta''} u^{\gamma''} u^{\delta''} - 2m \int_u^v V^{\alpha\beta}_{\ \ \ \mu\nu}(x, z) u^\mu u^\nu \, d\tau, \tag{19.60}$$

where primed indices refer to the retarded point $x' := z(u)$, double-primed indices refer to the advanced point $x'' := z(v)$, and where $r_{\rm adv} := -\sigma_{\alpha''} u^{\alpha''}$ is the advanced distance between x and the world line.

Differentiation of Eq. (19.60) yields

$$\begin{aligned} \gamma^{\rm S}_{\alpha\beta;\gamma}(x) \ = \ & -\frac{2m}{r^2} U_{\alpha\beta\alpha'\beta'} u^{\alpha'} u^{\beta'} \partial_\gamma r - \frac{2m}{{r_{\rm adv}}^2} U_{\alpha\beta\alpha''\beta''} u^{\alpha''} u^{\beta''} \partial_\gamma r_{\rm adv} + \frac{2m}{r} U_{\alpha\beta\alpha'\beta';\gamma} u^{\alpha'} u^{\beta'} \\ & + \frac{2m}{r} U_{\alpha\beta\alpha'\beta';\gamma'} u^{\alpha'} u^{\beta'} u^{\gamma'} \partial_\gamma u + \frac{2m}{r_{\rm adv}} U_{\alpha\beta\alpha''\beta'';\gamma} u^{\alpha''} u^{\beta''} + \frac{2m}{r_{\rm adv}} U_{\alpha\beta\alpha''\beta'';\gamma''} u^{\alpha''} u^{\beta''} u^{\gamma''} \partial_\gamma v \\ & + 2m V_{\alpha\beta\alpha'\beta'} u^{\alpha'} u^{\beta'} \partial_\gamma u - 2m V_{\alpha\beta\alpha''\beta''} u^{\alpha''} u^{\beta''} \partial_\gamma v - 2m \int_u^v \nabla_\gamma V_{\alpha\beta\mu\nu}(x, z) u^\mu u^\nu \, d\tau, \end{aligned} \tag{19.61}$$

and we would like to express this as an expansion in powers of r. For this we will rely on results already established in Sec. 19.3, as well as additional expansions that will involve the advanced point x''. We recall that a relation between retarded and advanced times was worked out in Eq. (11.12), that an expression for the advanced distance was displayed in Eq. (11.13), and that Eqs. (11.14) and (11.15) give expansions for $\partial_\gamma v$ and $\partial_\gamma r_{\rm adv}$, respectively; these results can be simplified by setting $a_a = \dot{a}_0 = \dot{a}_a = 0$, which is appropriate in this computation.

To derive an expansion for $U_{\alpha\beta\alpha''\beta''} u^{\alpha''} u^{\beta''}$ we follow the general method of Sec. 11.4 and introduce the functions $U_{\alpha\beta}(\tau) := U_{\alpha\beta\mu\nu}(x, z) u^\mu u^\nu$. We have that

$$U_{\alpha\beta\alpha''\beta''} u^{\alpha''} u^{\beta''} := U_{\alpha\beta}(v) = U_{\alpha\beta}(u) + \dot{U}_{\alpha\beta}(u)\Delta' + \frac{1}{2}\ddot{U}_{\alpha\beta}(u)\Delta'^2 + O\bigl(\Delta'^3\bigr),$$

where overdots indicate differentiation with respect to τ and $\Delta' := v - u$. The leading term $U_{\alpha\beta}(u) := U_{\alpha\beta\alpha'\beta'} u^{\alpha'} u^{\beta'}$ was worked out in Eq. (19.35), and the derivatives of $U_{\alpha\beta}(\tau)$ are given by

$$\dot{U}_{\alpha\beta}(u) = U_{\alpha\beta\alpha'\beta';\gamma'} u^{\alpha'} u^{\beta'} u^{\gamma'} = g^{\alpha'}_{\ (\alpha} g^{\beta'}_{\ \beta)} \Bigl[r R_{\alpha'0d0} \Omega^d u_{\beta'} + O(r^2) \Bigr]$$

and

$$\ddot{U}_{\alpha\beta}(u) = U_{\alpha\beta\alpha'\beta';\gamma'\delta'} u^{\alpha'} u^{\beta'} u^{\gamma'} u^{\delta'} = O(r),$$

according to Eqs. (19.37) and (16.15). Combining these results together with Eq. (11.12) for Δ' gives

$$U_{\alpha\beta\alpha''\beta''} u^{\alpha''} u^{\beta''} = g^{\alpha'}_{\ (\alpha} g^{\beta'}_{\ \beta)} \Bigl[u_{\alpha'} u_{\beta'} + 2r^2 R_{\alpha'0d0} \Omega^d u_{\beta'} + O(r^3) \Bigr], \tag{19.62}$$

which should be compared with Eq. (19.35). It should be emphasized that in Eq. (19.62) and all equations below, all frame components are evaluated at the retarded point x', and not at the advanced point. The preceding computation gives us also an expansion for

$$U_{\alpha\beta\alpha''\beta'';\gamma''} u^{\alpha'} u^{\beta''} u^{\gamma''} = \dot{U}_{\alpha\beta}(u) + \ddot{U}_{\alpha\beta}(u)\Delta' + O(\Delta'^2),$$

which becomes

$$U_{\alpha\beta\alpha''\beta'';\gamma''} u^{\alpha''} u^{\beta''} u^{\gamma''} = g^{\alpha'}_{\ (\alpha} g^{\beta'}_{\ \beta)} \Bigl[r R_{\alpha'0d0} \Omega^d u_{\beta'} + O(r^2) \Bigr], \tag{19.63}$$

and which is identical to Eq. (19.37).

We proceed similarly to obtain an expansion for $U_{\alpha\beta\alpha''\beta'';\gamma} u^{\alpha''} u^{\beta''}$. Here we introduce the functions $U_{\alpha\beta\gamma}(\tau) := U_{\alpha\beta\mu\nu;\gamma} u^\mu u^\nu$ and express $U_{\alpha\beta\alpha''\beta'';\gamma} u^{\alpha''} u^{\beta''}$ as $U_{\alpha\beta\gamma}(v) = U_{\alpha\beta\gamma}(u) + \dot{U}_{\alpha\beta\gamma}(u)\Delta' + O(\Delta'^2)$. The leading term $U_{\alpha\beta\gamma}(u) := U_{\alpha\beta\alpha'\beta';\gamma} u^{\alpha'} u^{\beta'}$ was computed in Eq. (19.36), and

$$\dot{U}_{\alpha\beta\gamma}(u) = U_{\alpha\beta\alpha'\beta';\gamma\gamma'} u^{\alpha'} u^{\beta'} u^{\gamma'} = g^{\alpha'}_{\ (\alpha} g^{\beta'}_{\ \beta)} g^{\gamma'}_{\ \gamma} \Bigl[R_{\alpha'0\gamma'0} u_{\beta'} + O(r) \Bigr]$$

follows from Eq. (16.14). Combining these results together with Eq. (11.12) for Δ' gives

$$U_{\alpha\beta\alpha''\beta'';\gamma} u^{\alpha''} u^{\beta''} = g^{\alpha'}_{\ (\alpha} g^{\beta'}_{\ \beta)} g^{\gamma'}_{\ \gamma} \Bigl[r \bigl(R_{\alpha'0\gamma'0} - R_{\alpha'0\gamma'd} \Omega^d \bigr) u_{\beta'} + O(r^2) \Bigr], \tag{19.64}$$

and this should be compared with Eq. (19.36). The last expansion we shall need is

$$V_{\alpha\beta\alpha''\beta''} u^{\alpha''} u^{\beta''} = g^{\alpha'}_{\ (\alpha} g^{\beta'}_{\ \beta)} \Bigl[R_{\alpha'0\beta'0} + O(r) \Bigr], \tag{19.65}$$

which is identical to Eq. (19.38).

We obtain the frame components of the singular gravitational field by substituting these expansions into Eq. (19.61) and projecting against the tetrad (e^α_0, e^α_a). After some algebra we arrive at

$$\begin{aligned}
\gamma^{\rm S}_{000}(u,r,\Omega^a) &:= \gamma^{\rm S}_{\alpha\beta;\gamma}(x) e^\alpha_0(x) e^\beta_0(x) e^\gamma_0(x) = 2m R_{a0b0} \Omega^a \Omega^b + O(r), &(19.66)\\
\gamma^{\rm S}_{0b0}(u,r,\Omega^a) &:= \gamma^{\rm S}_{\alpha\beta;\gamma}(x) e^\alpha_0(x) e^\beta_b(x) e^\gamma_0(x) = -4m R_{b0c0} \Omega^c + O(r), &(19.67)\\
\gamma^{\rm S}_{ab0}(u,r,\Omega^a) &:= \gamma^{\rm S}_{\alpha\beta;\gamma}(x) e^\alpha_a(x) e^\beta_b(x) e^\gamma_0(x) = O(r), &(19.68)\\
\gamma^{\rm S}_{00c}(u,r,\Omega^a) &:= \gamma^{\rm S}_{\alpha\beta;\gamma}(x) e^\alpha_0(x) e^\beta_0(x) e^\gamma_c(x) \\
&= -4m \biggl[\Bigl(\frac{1}{r^2} + \frac{1}{3} R_{a0b0} \Omega^a \Omega^b \Bigr) \Omega_c + \frac{1}{6} R_{c0b0} \Omega^b - \frac{1}{6} R_{ca0b} \Omega^a \Omega^b \biggr] + O(r), &(19.69)\\
\gamma^{\rm S}_{0bc}(u,r,\Omega^a) &:= \gamma^{\rm S}_{\alpha\beta;\gamma}(x) e^\alpha_0(x) e^\beta_b(x) e^\gamma_c(x) = 2m \bigl(R_{b0cd} \Omega^d + R_{b0d0} \Omega^d \Omega_c \bigr) + O(r), &(19.70)\\
\gamma^{\rm S}_{abc}(u,r,\Omega^a) &:= \gamma^{\rm S}_{\alpha\beta;\gamma}(x) e^\alpha_a(x) e^\beta_b(x) e^\gamma_c(x) = -4m R_{a0b0} \Omega_c + O(r), &(19.71)
\end{aligned}$$

in which all frame components are evaluated at the retarded point x'. Comparison of these expressions with Eqs. (19.39)–(19.44) reveals identical singularity structures for the retarded and singular gravitational fields.

The difference between the retarded field of Eqs. (19.39)–(19.44) and the singular field of Eqs. (19.66)–(19.71) defines the regular gravitational field $\gamma^{\rm R}_{\alpha\beta;\gamma}$. Its frame components are

$$\begin{aligned}
\gamma^{\rm R}_{000} &= \gamma^{\rm tail}_{000} + O(r), &(19.72)\\
\gamma^{\rm R}_{0b0} &= \gamma^{\rm tail}_{0b0} + O(r), &(19.73)\\
\gamma^{\rm R}_{ab0} &= 4m R_{a0b0} + \gamma^{\rm tail}_{ab0} + O(r), &(19.74)\\
\gamma^{\rm R}_{00c} &= \gamma^{\rm tail}_{00c} + O(r), &(19.75)\\
\gamma^{\rm R}_{0bc} &= 2m R_{b0c0} + \gamma^{\rm tail}_{0bc} + O(r), &(19.76)\\
\gamma^{\rm R}_{abc} &= \gamma^{\rm tail}_{abc} + O(r), &(19.77)
\end{aligned}$$

and we see that $\gamma^{\rm R}_{\alpha\beta;\gamma}$ is regular in the limit $r \to 0$. We may therefore evaluate the regular field directly at $x = x'$, where the tetrad (e^α_0, e^α_a) coincides with $(u^{\alpha'}, e^{\alpha'}_a)$. After reconstructing the field at x' from its frame components, we obtain

$$\gamma^{\rm R}_{\alpha'\beta';\gamma'}(x') = -4m\Big(u_{(\alpha'}R_{\beta')\delta'\gamma'\epsilon'} + R_{\alpha'\delta'\beta'\epsilon'}u_{\gamma'}\Big)u^{\delta'}u^{\epsilon'} + \gamma^{\rm tail}_{\alpha'\beta'\gamma'}, \tag{19.78}$$

where the tail term can be copied from Eq. (19.34),

$$\gamma^{\rm tail}_{\alpha'\beta'\gamma'}(x') = 4m\int_{-\infty}^{u^-} \nabla_{\gamma'}G_{+\alpha'\beta'\mu\nu}(x', z)u^\mu u^\nu\, d\tau. \tag{19.79}$$

The tensors that appear in Eq. (19.79) all refer to the retarded point $x' := z(u)$, which can now be treated as an arbitrary point on the world line γ.

19.6 Equations of motion

The retarded gravitational field $\gamma_{\alpha\beta;\gamma}$ of a point particle is singular on the world line, and this behaviour makes it difficult to understand how the field is supposed to act on the particle and influence its motion. The field's singularity structure was analyzed in Secs. 19.3 and 19.4, and in Sec. 19.5 it was shown to originate from the singular field $\gamma^{\rm S}_{\alpha\beta;\gamma}$; the regular field $\gamma^{\rm R}_{\alpha\beta;\gamma}$ was then shown to be regular on the world line.

To make sense of the retarded field's action on the particle we can follow the discussions of Sec. 17.6 and 18.6 and postulate that the self gravitational field of the point particle is either $\langle\gamma_{\mu\nu;\lambda}\rangle$, as worked out in Eq. (19.57), or $\gamma^{\rm R}_{\mu\nu;\lambda}$, as worked out in Eq. (19.78). These regularized fields are both given by

$$\gamma^{\rm reg}_{\mu\nu;\lambda} = -4m\Big(u_{(\mu}R_{\nu)\rho\lambda\xi} + R_{\mu\rho\nu\xi}u_\lambda\Big)u^\rho u^\xi + \gamma^{\rm tail}_{\mu\nu\lambda} \tag{19.80}$$

and

$$\gamma^{\rm tail}_{\mu\nu\lambda} = 4m\int_{-\infty}^{\tau^-} \nabla_\lambda G_{+\mu\nu\mu'\nu'}\big(z(\tau), z(\tau')\big)u^{\mu'}u^{\nu'}\, d\tau', \tag{19.81}$$

in which all tensors are now evaluated at an arbitrary point $z(\tau)$ on the world line γ.

The actual gravitational perturbation $h_{\alpha\beta}$ is obtained by inverting Eq. (19.10), which leads to $h_{\mu\nu;\lambda} = \gamma_{\mu\nu;\gamma} - \frac{1}{2}g_{\mu\nu}\gamma^\rho{}_{\rho;\lambda}$. Substituting Eq. (19.80) yields

$$h^{\rm reg}_{\mu\nu;\lambda} = -4m\Big(u_{(\mu}R_{\nu)\rho\lambda\xi} + R_{\mu\rho\nu\xi}u_\lambda\Big)u^\rho u^\xi + h^{\rm tail}_{\mu\nu\lambda}, \tag{19.82}$$

where the tail term is given by the trace-reversed counterpart to Eq. (19.81):

$$h^{\rm tail}_{\mu\nu\lambda} = 4m\int_{-\infty}^{\tau^-} \nabla_\lambda\Big(G_{+\mu\nu\mu'\nu'} - \frac{1}{2}g_{\mu\nu}G_+{}^\rho{}_{\rho\mu'\nu'}\Big)\big(z(\tau), z(\tau')\big)u^{\mu'}u^{\nu'}\, d\tau'. \tag{19.83}$$

When this regularized field is substituted into Eq. (19.13), we find that the terms that depend on the Riemann tensor cancel out, and we are left with

$$\frac{Du^\mu}{d\tau} = -\frac{1}{2}\big(g^{\mu\nu} + u^\mu u^\nu\big)\big(2h^{\rm tail}_{\nu\lambda\rho} - h^{\rm tail}_{\lambda\rho\nu}\big)u^\lambda u^\rho. \tag{19.84}$$

We see that only the tail term is involved in the final form of the equations of motion. The tail integral of Eq. (19.83) involves the current position $z(\tau)$ of the particle, at which all tensors with unprimed indices are evaluated, as well as all prior positions $z(\tau')$, at which all tensors with primed indices are evaluated. The tail integral is cut short at $\tau' = \tau^- := \tau - 0^+$ to avoid the singular behaviour of the retarded Green's function at coincidence; this limiting procedure was justified at the beginning of Sec. 19.3.

Equation (19.84) was first derived by Yasushi Mino, Misao Sasaki, and Takahiro Tanaka in 1997 [6]. (An incomplete treatment had been given previously by Morette-DeWitt and Ging [151].) An alternative derivation was then produced, also in 1997, by Theodore C. Quinn and Robert M. Wald [7]. These equations

are now known as the MiSaTaQuWa equations of motion, and other derivations [16,17], based on an extended-body approach, will be reviewed below in Part V. It should be noted that Eq. (19.84) is formally equivalent to the statement that the point particle moves on a geodesic in a spacetime with metric $g_{\alpha\beta} + h^{\rm R}_{\alpha\beta}$, where $h^{\rm R}_{\alpha\beta}$ is the regular metric perturbation obtained by trace-reversal of the potentials $\gamma^{\rm R}_{\alpha\beta} := \gamma_{\alpha\beta} - \gamma^{\rm S}_{\alpha\beta}$; this perturbed metric is regular on the world line, and it is a solution to the vacuum field equations. This elegant interpretation of the MiSaTaQuWa equations was proposed in 2003 by Steven Detweiler and Bernard F. Whiting [13]. Quinn and Wald [141] have shown that under some conditions, the total work done by the gravitational self-force is equal to the energy radiated (in gravitational waves) by the particle.

Part V
Motion of a small body

20 Point-particle limits and matched asymptotic expansions

The expansion presented in the previous section is based on an exact point-particle source. But in the full, nonlinear theory, no distributional solution would exist for such a source [142]. Although the expansion nevertheless yields a well-behaved linear approximation, it is ill-behaved beyond that order, since the second- and higher-order Einstein tensors will contain products of delta functions, which have no meaning as distributions. It may be possible to overcome this limitation using more advanced methods such as Colombeau algebras [152], which allow for the multiplication of distributions, but little work has been done to that end. Instead, the common approach, and the one we shall pursue here, has been to abandon the fiction of a point particle in favor of considering an asymptotically small body. As we shall see, we can readily generalize the self-consistent expansion scheme to this case. Furthermore, we shall find that the results of the previous section are justified by this approach: at linear order, the metric perturbation due to an asymptotically small body is precisely that of a point particle moving on a world line with an acceleration given by the MiSaTaQuWa equation (plus higher-order corrections).

In order for the body to be considered "small," its mass and size must be much smaller than all external lengthscales. We denote these external scales collectively as $\mathscr{R}$, which we may define to be the radius of curvature of the spacetime (were the small body removed from it) in the region in which we seek an approximation. Given this definition, a typical component of the spacetime's Riemann tensor is equal to $1/\mathscr{R}^2$ up to a numerical factor of order unity. Now, we consider a family of metrics $g_{\alpha\beta}(\varepsilon)$ containing a body whose mass scales as ε in the limit $\varepsilon \to 0$; that is, $\varepsilon \sim m/\mathscr{R}$. If each member of the family is to contain a body of the same type, then the size of the body must also approach zero with ε. The precise scaling of size with ε is determined by the type of body, but it is not generally relevant. What *is* relevant is the "gravitational size" — the length scale determining the metric outside the body — and this size always scales linearly with the mass. If the body is compact, as is a neutron star or a black hole, then its gravitational size is also its actual linear size. In what remains, we assume that all lengths have been scaled by $\mathscr{R}$, such that we can write, for example, $m \ll 1$. Our goal is to determine the metric perturbation and the equation of motion produced by the body in this limit.

Point-particle limits such as this have been used to derive equations of motion many times in the past, including in derivations of geodesic motion at leading order [153–155] and in constructing post-Newtonian limits [156]. Perhaps the most obvious means of approaching the problem is to first work nonperturbatively, with a body of arbitrary size, and then take the limit. Using this approach (though with some restrictions on the body's size and compactness) and generalized definitions of momenta, Harte has calculated the self-force in the case of scalar [157] and electromagnetic [158] charge distributions in fixed backgrounds, following the earlier work of Dixon [159–161]. However, while this approach is conceptually compelling, at this stage it applies only to material bodies, not black holes, and has not yet been presented as part of a systematic expansion of the Einstein equation. Here, we focus instead on a more general method.

Alternatively, one could take the opposite approach, essentially taking the limit first and then trying to recover the higher-order effects, by treating the body as an *effective* point particle at leading order, with finite size effects introduced as higher-order effective fields, as done by Galley and Hu [162, 163]. However, while this approach is computationally efficient, allowing one to perform high-order calculations with (relative) ease, it requires methods such as dimensional regularization and mass renormalization in order to arrive at meaningful results. Because of these undesirable requirements, we will not consider it here.

In the approach we review, we make use of the method of matched asymptotic expansions [2, 6, 15–18, 33, 138, 156, 164–170]. Broadly speaking, this method consists of constructing two different asymptotic expansions, each valid in a specific region, and combining them to form a global expansion. In the present context, the method begins with two types of point-particle limits: an *outer limit*, in which $\varepsilon \to 0$ at fixed

coordinate values (we will slightly modify this in a moment); and an *inner limit*, in which $\varepsilon \to 0$ at fixed values of $\tilde{R} := R/\varepsilon$, where R is a measure of radial distance from the body. In the outer limit, the body shrinks toward zero size as all other distances remain roughly constant; in the inner limit, the small body keeps a constant size while all other distances blow up toward infinity. Thus, the inner limit serves to "zoom in" on a small region around the body. The outer limit can be expected to be valid in regions where $R \sim 1$, while the inner limit can be expected to be valid in regions where $\tilde{R} \sim 1$ (or $R \sim \varepsilon$), though both of these regions can be extended into larger domains.

More precisely, consider an exact solution $\mathsf{g}_{\alpha\beta}$ on a manifold $\mathcal{M}_\varepsilon$ with two coordinate systems: a local coordinate system $X^\alpha = (T, R, \Theta^A)$ that is centered (in some approximate sense) on the small body, and a global coordinate system x^α. For example, in an extreme-mass-ratio inspiral, the local coordinates might be the Schwarzschild-type coordinates of the small body, and the global coordinates might be the Boyer-Lindquist coordinates of the supermassive Kerr black hole. In the outer limit, we expand $\mathsf{g}_{\alpha\beta}$ for small ε while holding x^α fixed. The leading-order solution in this case is the background metric $g_{\alpha\beta}$ on a manifold $\mathcal{M}_E$; this is the external spacetime, which contains no small body. It might, for example, be the spacetime of the supermassive black hole. In the inner limit, we expand $\mathsf{g}_{\alpha\beta}$ for small ε while holding $(T, \tilde{R}, \Theta^A)$ fixed. The leading-order solution in this case is the metric $g^{\text{body}}_{\alpha\beta}$ on a manifold $\mathcal{M}_I$; this is the spacetime of the small body if it were isolated (though it may include slow evolution due to its interaction with the external spacetime — this will be discussed below). Note that $\mathcal{M}_E$ and $\mathcal{M}_I$ generically differ: in an extreme-mass-ratio inspiral, for example, if the small body is a black hole, then $\mathcal{M}_I$ will contain a spacelike singularity in the black hole's interior, while $\mathcal{M}_E$ will be smooth at the "position" where the small black hole would be. What we are interested in is that "position" — the world line in the smooth external spacetime $\mathcal{M}_E$ that represents the motion of the small body. Note that this world line generically appears *only* in the external spacetime, rather than as a curve in the exact spacetime $(\mathsf{g}_{\alpha\beta}, \mathcal{M}_\varepsilon)$; in fact, if the small body is a black hole, then obviously no such curve exists.

Determining this world line presents a fundamental problem. In the outer limit, the body vanishes at $\varepsilon = 0$, leaving only a remnant, ε-independent curve in $\mathcal{M}_E$. (Outside any small body, the metric will contain terms such as m/R, such that in the limit $m \to 0$, the limit exists everywhere except at $R = 0$, which leaves a removable discontinuity in the external spacetime; the removal of this discontinuity defines the remnant world line of the small body.) But the true motion of the body will generically be ε-dependent. If we begin with the remnant world line and correct it with the effects of the self-force, for example, then the corrections must be small: they are small deviation vectors defined on the remnant world line. Put another way, if we expand $\mathsf{g}_{\alpha\beta}$ in powers of ε, then all functions in it must similarly be expanded, including any representation of the motion, and in particular, any representative world line. We would then have a representation of the form $z^\alpha(t, \varepsilon) = z^\alpha_{(0)}(t) + \varepsilon z^\alpha_{(1)}(t) + \ldots$, where $z^\alpha_{(1)}(t)$ is a vector defined on the remnant curve described by $z^\alpha_{(0)}(t)$. The remnant curve would be a geodesic, and the small corrections would incorporate the self-force and finite-size effects [16] (see also [166]). However, because the body will generically drift away from any such geodesic, the small corrections will generically grow large with time, leading to the failure of the regular expansion. So we will modify this approach by performing a self-consistent expansion in the outer limit, following the same scheme as presented in the point-particle case. Refs. [17,18,129] contain far more detailed discussions of these points.

Regardless of whether the self-consistent expansion is used, the success of matched asymptotic expansions relies on the buffer region defined by $\varepsilon \ll R \ll 1$ (see Fig. 10). In this region, both the inner and outer expansions are valid. From the perspective of the outer expansion, this corresponds to an asymptotically small region around the world line: $R \ll 1$. From the perspective of the inner expansion, it corresponds to asymptotic spatial infinity: $1/\tilde{R} = \varepsilon/R \ll 1$. Because both expansions are valid in this region, and because both are expansions of the same exact metric $\mathsf{g}_{\alpha\beta}$ and hence must "match," by working in the buffer region we can use information from the inner expansion to determine information about the outer expansion (or vice versa). We shall begin by solving the Einstein equation in the buffer region, using information from the inner expansion to determine the form of the external metric perturbation therein. In so doing, we shall determine the acceleration of the small body's world line. Finally, using the field values in the buffer region, we shall construct a global solution for the metric perturbation.

In this calculation, the structure of the body is left unspecified. Our only condition is that part of the buffer region must lie outside the body, because we wish to solve the Einstein field equations in vacuum.

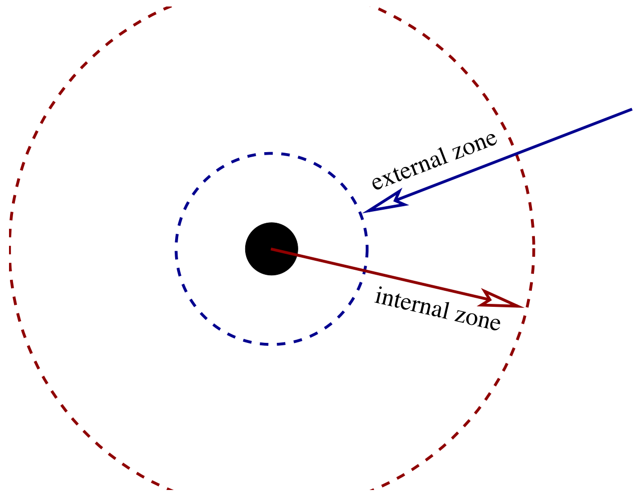

Figure 10: A small body, represented by the black disk, is immersed in a background spacetime. The internal zone is defined by $R \ll 1$, while the external zone is defined by $R \gg \varepsilon$. Since $\varepsilon \ll 1$, there exists a buffer region defined by $\varepsilon \ll R \ll 1$. In the buffer region ε/R and R are both small.

This requires the body to be sufficiently compact. For example, our calculation would fail for a diffuse body such as our Sun; likewise, it would fail if a body became tidally disrupted. Although we will detail only the case of an uncharged body, the same techniques would apply to charged bodies; Gralla *et al.* [2] have recently performed a similar calculation for the electromagnetic self-force on an asymptotically small body in a flat background spacetime. Using very different methods, Futamase *et al.* [171] have calculated equations of motion for an asymptotically small charged black hole.

The structure of our discussion is as follows: In Sec. 21, we present the self-consistent expansion of the Einstein equation. Next, in Sec. 22, we solve the equations in the buffer region up to second order in the outer expansion. Last, in Sec. 23, we discuss the global solution in the outer expansion and show that it is that of a point particle at first order. Over the course of this calculation, we will take the opportunity to incorporate several details that we could have accounted for in the point-particle case but opted to neglect for simplicity: an explicit expansion of the acceleration vector that makes the self-consistent expansion properly systematic, and a finite time domain that accounts for the fact that large errors eventually accumulate if the approximation is truncated at any finite order. For more formal discussions of matched asymptotic expansions in general relativity, see Refs. [18, 172]; the latter reference, in particular, discusses the method as it pertains to the motion of small bodies. For background on the use of matched asymptotic expansions in applied mathematics, see Refs. [131, 145–148]; the text by Eckhaus [145] provides the most rigorous treatment.

21 Self-consistent expansion

21.1 Introduction

We wish to represent the motion of the body through the external background spacetime $(g_{\alpha\beta}, \mathcal{M}_E)$, rather than through the exact spacetime $(\mathsf{g}_{\alpha\beta}, \mathcal{M}_\varepsilon)$. In order to achieve this, we begin by surrounding the body with a (hollow, three-dimensional) world tube Γ embedded in the buffer region. We define the tube to be a surface of constant radius $s = \mathcal{R}(\varepsilon)$ in Fermi normal coordinates centered on a world line $\gamma \subset \mathcal{M}_E$, though the exact definition of the tube is immaterial. Since there exists a diffeomorphism between $\mathcal{M}_E$ and $\mathcal{M}_I$ in the buffer region, this defines a tube $\Gamma_I \subset \mathcal{M}_I$. Now, the problem is the following: what equation of motion must γ satisfy in order for Γ_I to be "centered" about the body?

How shall we determine if the body lies at the centre of the tube's interior? Since the tube is close to the small body (relative to all external length scales), the metric on the tube is primarily determined by the small body's structure. Recall that the buffer region corresponds to an asymptotically large spatial distance

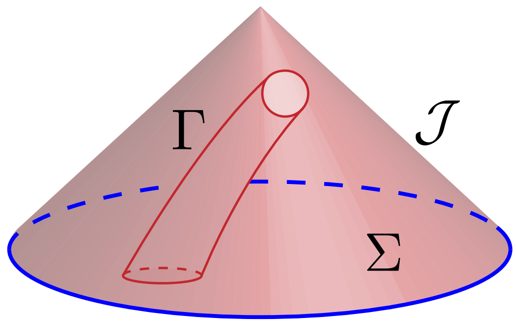

Figure 11: The spacetime region Ω is bounded by the union of the spacelike surface Σ, the timelike tube Γ, and the null surface $\mathcal{J}$.

in the inner expansion. Hence, on the tube, we can construct a multipole expansion of the body's field, with the form $\sum R^{-n}$ (or $\sum s^{-n}$—we will assume $s \sim R$ in the buffer region). Although alternative definitions could be used, we define the tube to be centered about the body if the mass dipole moment vanishes in this expansion. Note that this is the typical approach in general relativity: Whereas in Newtonian mechanics one directly finds the equation of motion for the centre of mass of a body, in general relativity one typically seeks a world line about which the mass dipole of the body vanishes (or an equation of motion for the mass dipole relative to a given nearby world line) [16, 17, 173, 174]. This definition of the world line is sufficiently general to apply to a black hole. If the body is material, one could instead imagine a centre-of-mass world line that lies in the interior of the body in the exact spacetime. This world line would then be the basis of our self-consistent expansion. We use our more general definition to cover both cases. See Ref. [175] and references therein for discussion of multipole expansions in general relativity, see Refs. [138, 175] for discussions of mass-centered coordinates in the buffer region, and see, e.g., Refs. [176, 177] for alternative definitions of centre of mass in general relativity.

As in the point-particle case, in order to determine the equation of motion of the world line, we consider a family of metrics, now denoted $g_E(x,\varepsilon;\gamma)$, parametrized by γ, such that when γ is given by the correct equation of motion for a given value of ε, we have $g_E(x,\varepsilon;\gamma(\varepsilon)) = \mathsf{g}(x)$. The metric in the outer limit is thus taken to be the general expansion

$$\mathsf{g}_{\alpha\beta}(x,\varepsilon) = g_{E\alpha\beta}(x,\varepsilon;\gamma) = g_{\alpha\beta}(x) + h_{\alpha\beta}(x,\varepsilon;\gamma), \tag{21.1}$$

where

$$h_{\alpha\beta}(x,\varepsilon;\gamma) = \sum_{n=1}^{\infty} \varepsilon^n h^{(n)}_{\alpha\beta}(x;\gamma). \tag{21.2}$$

In the point-particle case, solving Einstein's equations determined the equation of motion of the particle's world line; in this case, it will determine the world line γ for which the inner expansion is mass-centered. In this self-consistent expansion, the perturbations produced by the body are constructed about a fixed world line determined by the particular value of ε at which one seeks an approximation.

In the remainder of this section, we present a sequence of perturbation equations that arise in this expansion scheme, along with a complementary sequence for the inner expansion.

21.2 Field equations in outer expansion

In the outer expansion, we seek a solution in a vacuum region Ω outside of Γ. We specify $\Omega \subset \mathcal{M}_E$ to be an open set consisting of the future domain of dependence of the spacelike initial-data surface Σ, excluding the interior of the world tube Γ. This implies that the future boundary of Ω is a null surface $\mathcal{J}$. Refer to

Fig. 11 for an illustration. The boundary of the domain is $\partial\Omega := \Gamma \cup \mathcal{J} \cup \Sigma$. The spatial surface Σ is chosen to intersect Γ at the initial time $t = 0$.

Historically, in derivations of the self-force, solutions to the perturbative field equations were taken to be global in time, with tail integrals extended to negative infinity, as we wrote them in the preceding sections. But as was first noted in Ref. [17], because the self-force drives long-term, cumulative changes, any approximation truncated at a given order will be accurate to that order only for a finite time; and this necessites working in a finite region such as Ω. This is also true in the case of point charges and masses. For simplicity, we neglected this detail in the preceding sections, but for completeness, we account for it here.

Field equations

Within this region, we follow the methods presented in the case of a point mass. We begin by reformulating the Einstein equation such that it can be solved for an arbitrary world line. To accomplish this, we assume that the Lorenz gauge can be imposed on the whole of $h_{\alpha\beta}$, everywhere in Ω, such that $L_\mu[h] = 0$. Here

$$L_\mu[h] := \left(g^{\alpha}_{\ \mu} g^{\beta\nu} - \frac{1}{2} g^{\alpha\beta} g_{\mu}^{\ \nu}\right) \nabla_\nu h_{\alpha\beta} \tag{21.3}$$

is the Lorenz-gauge operator that was first introduced in Secs. 16.1 and 19.1; the condition $L_\mu[h] = 0$ is the same statement as $\nabla_\nu \gamma^{\mu\nu} = 0$, where $\gamma_{\mu\nu} := h_{\mu\nu} - \frac{1}{2} g_{\mu\nu} g^{\alpha\beta} h_{\alpha\beta}$ is the "trace-reversed" metric perturbation. We discuss the validity of this assumption below.

Just as in the case of a point mass, this choice of gauge reduces the vacuum Einstein equation $\mathsf{R}_{\mu\nu} = 0$ to a weakly nonlinear wave equation that can be expanded and solved at fixed γ. However, we now seek a solution only in the region Ω, where the energy-momentum tensor vanishes, so the resulting sequence of wave equations reads

$$E_{\mu\nu}[h^{(1)}] = 0, \tag{21.4}$$

$$E_{\mu\nu}[h^{(2)}] = 2\delta^2 R_{\mu\nu}[h^{(1)}], \tag{21.5}$$

$$\vdots$$

where

$$E_{\mu\nu}[h] := \left(g^{\alpha}_{\ \mu} g^{\beta}_{\ \nu} \nabla^\gamma \nabla_\gamma + 2R_{\mu\ \nu}^{\ \alpha\ \beta}\right) h_{\alpha\beta} \tag{21.6}$$

is the tensorial wave operator that was first introduced in Secs. 16.1 and 19.1, and the second-order Ricci tensor $\delta^2 R_{\mu\nu}$, which is quadratic in its argument, is shown explicitly in Eq (A.1). More generally, we can write the nth-order equation as

$$E_{\mu\nu}\big[h^{(n)}\big] = S^{(n)}_{\mu\nu}\big[h^{(1)}, \dots, h^{(n-1)}, \gamma\big], \tag{21.7}$$

where the source term $S^{(n)}_{\mu\nu}$ consists of nonlinear terms in the expansion of the Ricci tensor.

Again as in the case of a point particle, we can easily write down formal solutions to the wave equations, for arbitrary γ. Using the same methods as were used to derive the Kirchoff representation in Sec. 16.3, we find

$$h^{(n)}_{\alpha\beta} = \frac{1}{4\pi} \oint_{\partial\Omega} \left(G^{+\ \gamma'\delta'}_{\alpha\beta} \nabla_{\mu'} h^{(n)}_{\gamma'\delta'} - h^{(n)}_{\gamma'\delta'} \nabla_{\mu'} G^{+\ \gamma'\delta'}_{\alpha\beta}\right) dS^{\mu'} + \frac{1}{4\pi} \int_\Omega G^{+\ \gamma'\delta'}_{\alpha\beta} S^{(n)}_{\gamma'\delta'} dV'. \tag{21.8}$$

Because $\mathcal{J}$ is a future null surface, the integral over it vanishes. Hence, this formal solution requires only initial data on Σ and boundary data on Γ. Since Γ lies in the buffer region, the boundary data on it is determined by information from the inner expansion.

One should note several important properties of these integral representations: First, x must lie in the interior of Ω; an alternative expression must be derived if x lies on the boundary [178]. Second, the integral over the boundary is, in each case, a homogeneous solution to the wave equation, while the integral over the volume is an inhomogeneous solution. Third, if the field at the boundary satisfies the Lorenz gauge condition,

then by virtue of the wave equation, it satisfies the gauge condition everywhere; hence, imposing the gauge condition to some order in the buffer region ensures that it is imposed to the same order everywhere.

While the integral representation is satisfied by any solution to the associated wave equation, it does not *provide* a solution. That is, one cannot prescribe arbitrary boundary values on Γ and then arrive at a solution. The reason is that the tube is a timelike boundary, which means that field data on it can propagate forward in time and interfere with the data at a later time. However, by applying the wave operator $E_{\alpha\beta}$ onto equation (21.8), we see that the integral representation of $h^{(n)}_{\alpha\beta}$ is guaranteed to satisfy the wave equation at each point $x \in \Omega$. In other words, the problem arises not in satisfying the wave equation in a pointwise sense, but in simultaneously satisfying the boundary conditions. But since the tube is chosen to lie in the buffer region, these boundary conditions can be supplied by the buffer-region expansion. And as we will discuss in Sec. 23, because of the asymptotic smallness of the tube, the pieces of the buffer-region expansion diverging as s^{-n} are sufficient boundary data to fully determine the global solution.

Finally, just as in the point-particle case, in order to split the gauge condition into a set of exactly solvable equations, we assume that the acceleration of γ possesses an expansion

$$a^\mu(t,\varepsilon) = a^\mu_{(0)}(t) + \varepsilon a^\mu_{(1)}(t;\gamma) + \dots. \tag{21.9}$$

This leads to the set of equations

$$L^{(0)}_\mu\big[h^{(1)}\big] = 0, \tag{21.10}$$
$$L^{(1)}_\mu\big[h^{(1)}\big] = -L^{(0)}_\mu\big[h^{(2)}\big], \tag{21.11}$$
$$\vdots$$

or, more generally, for $n > 0$,

$$L^{(n)}_\mu\big[h^{(1)}\big] = -\sum_{m=1}^{n} L^{(n-m)}_\mu\big[h^{(m+1)}\big]. \tag{21.12}$$

In these expressions, $L^{(0)}_\mu[f]$ is the Lorenz-gauge operator acting on the tensor field $f_{\alpha\beta}$ evaluated with $a^\mu = a^\mu_{(0)}$, $L^{(1)}_\mu[f]$ consists of the terms in $L_\mu[f]$ that are linear in $a^\mu_{(1)}$, and $L^{(n)}_\mu[f]$ contains the terms linear in $a^{(n)\mu}$, the combinations $a^{(n-1)\mu}a^{(1)\nu}$, and so on. Imposing these gauge conditions on the solutions to the wave equations will determine the acceleration of the world line. Although we introduce an expansion of the acceleration vector in order to obtain a systematic sequence of equations that can be solved exactly, such an expansion also trivially eliminates the need for order-reduction of the resulting equations of motion, since it automatically leads to equations for $a^{(1)\mu}$ in terms of $a^{(0)\mu}$, $a^{(2)\mu}$ in terms of $a^{(0)\mu}$ and $a^{(1)\mu}$, and so on.

Gauge transformations and the Lorenz condition

The outer expansion is defined not only by holding x^α fixed, but also by demanding that the mass dipole of the body vanishes when calculated in coordinates centered on γ. If we perform a gauge transformation generated by a vector $\xi^{(1)\alpha}(x;\gamma)$, then the mass dipole will no longer vanish in those coordinates. Hence, a new world line γ' must be constructed, such that the mass dipole vanishes when calculated in coordinates centered on that new world line. In other words, in the outer expansion we have the usual gauge freedom of regular perturbation theory, so long as the world line is appropriately transformed as well: $(h_{\alpha\beta},\gamma) \to (h'_{\alpha\beta},\gamma')$. The transformation law for the world line was first derived by Barack and Ori [19]; it was displayed in Eq. (1.49), and it will be worked out again in Sec. 22.6.

Using this gauge freedom, we now justify, to some extent, the assumption that the Lorenz gauge condition can be imposed on the entirety of $h_{\alpha\beta}$. If we begin with the metric in an arbitrary gauge, then the gauge vectors $\varepsilon\xi^\alpha_{(1)}[\gamma]$, $\varepsilon^2\xi^\alpha_{(2)}[\gamma]$, etc., induce the transformation

$$\begin{aligned} h_{\alpha\beta} \to h'_{\alpha\beta} &= h_{\alpha\beta} + \Delta h_{\alpha\beta} \\ &= h_{\alpha\beta} + \varepsilon \pounds_{\xi_{(1)}} g_{\alpha\beta} + \tfrac{1}{2}\varepsilon^2\big(\pounds_{\xi_{(2)}} + \pounds^2_{\xi_{(1)}}\big) g_{\alpha\beta} + \varepsilon^2 \pounds_{\xi_{(1)}} h^{(1)}_{\alpha\beta} + O(\varepsilon^3). \end{aligned} \tag{21.13}$$

If $h'_{\alpha\beta}$ is to satisfy the gauge condition $L_\mu[h']$, then ξ must satisfy $L_\mu[\Delta h] = -L_\mu[h]$. After a trivial calculation, this equation becomes

$$\sum_{n>0} \frac{\varepsilon^n}{n!}\Box\xi^\alpha_{(n)} = -\varepsilon L^\alpha\left[h^{(1)}\right] - \varepsilon^2 L^\alpha\left[h^{(2)}\right] - \varepsilon^2 L^\alpha\left[\tfrac{1}{2}\pounds^2_{\xi_{(1)}} g + \pounds_{\xi_{(1)}} h^{(1)}\right] + O(\varepsilon^3). \tag{21.14}$$

Solving this equation for arbitrary γ, we equate coefficients of powers of ε, leading to a sequence of wave equations of the form

$$\Box\xi^\alpha_{(n)} = W^\alpha_{(n)}, \tag{21.15}$$

where $W^\alpha_{(n)}$ is a functional of $\xi^\alpha_{(1)}, \ldots, \xi^\alpha_{(n-1)}$ and $h^{(1)}_{\alpha\beta}, \ldots, h^{(n)}_{\alpha\beta}$. We seek a solution in the region Ω described in the preceding section. The formal solution reads

$$\xi^\alpha_{(n)} = -\frac{1}{4\pi}\int_\Omega G^\alpha_{+\alpha'} W^{\alpha'}_{(n)} dV' + \frac{1}{4\pi}\oint_{\partial\Omega}\Big(G^\alpha_{+\gamma'}\nabla_{\mu'}\xi^{\gamma'}_{(n)} - \xi^{\gamma'}_{(n)}\nabla_{\mu'} G^\alpha_{+\gamma'}\Big) dS^{\mu'}. \tag{21.16}$$

From this we see that the Lorenz gauge condition can be adopted to any desired order of accuracy, given the existence of self-consistent data on a tube Γ of asymptotically small radius. We leave the question of the data's existence to future work. This argument was first presented in Ref. [18].

21.3 Field equations in inner expansion

For the inner expansion, we assume the existence of some local polar coordinates $X^\alpha = (T, R, \Theta^A)$, such that the metric can be expanded for $\varepsilon \to 0$ while holding fixed $\tilde{R} := R/\varepsilon$, Θ^A, and T; to relate the inner and outer expansions, we assume $R \sim s$, but otherwise leave the inner expansion completely general.

This leads to the ansatz

$$\mathsf{g}_{\alpha\beta}(T, \tilde{R}, \Theta^A, \varepsilon) = g^{\text{body}}_{\alpha\beta}(T, \tilde{R}, \Theta^A) + H_{\alpha\beta}(T, \tilde{R}, \Theta^A, \varepsilon), \tag{21.17}$$

where $H_{\alpha\beta}$ at fixed values of $(T, \tilde{R}, \Theta^A)$ is a perturbation beginning at order ε. This equation represents an asymptotic expansion along flow lines of constant R/ε as $\varepsilon \to 0$. It is tensorial in the usual sense of perturbation theory: the decomposition into $g^{\text{body}}_{\alpha\beta}$ and $H_{\alpha\beta}$ is valid in any coordinates that can be decomposed into ε-independent functions of the scaled coordinates plus $O(\varepsilon)$ functions of them. As written, with $g^{\text{body}}_{\alpha\beta}$ depending only on the scaled coordinates and independent of ε, the indices in Eq. (21.17) can be taken to refer to the unscaled coordinates (T, R, Θ^A). However, writing the components in the scaled coordinates will not alter the form of the expansion, but only introduce an overall rescaling of spatial components due to the spatial forms transforming as, e.g., $dR \to \varepsilon d\tilde{R}$. For example, if the body is a small Schwarzschild black hole of ADM mass $\varepsilon m(T)$, then in scaled Schwarzschild coordinates $(T, \tilde{R}, \Theta^A)$, $g^{\text{body}}_{\alpha\beta}(T, \tilde{R}, \Theta^A)$ is given by

$$ds^2 = -(1 - 2m(T)/\tilde{R})dT^2 + (1 - 2m(T)/\tilde{R})^{-1}\varepsilon^2 d\tilde{R}^2 + \varepsilon^2\tilde{R}^2(d\Theta^2 + \sin\Theta d\Phi^2). \tag{21.18}$$

As we would expect from the fact that the inner limit follows the body down as it shrinks, all points are mapped to the curve $R = 0$ at $\varepsilon = 0$, such that the metric in the scaled coordinates naturally becomes one-dimensional at $\varepsilon = 0$. This singular limit can be made regular by rescaling time as well, such that $\tilde{T} = (T - T_0)/\varepsilon$, and then rescaling the entire metric by a conformal factor $1/\varepsilon^2$. In order to arrive at a global-in-time inner expansion, rather than a different expansion at each time T_0, we forgo this extra step. We do, however, make an equivalent assumption, which is that the metric $g^{\text{body}}_{\alpha\beta}$ and its perturbations are quasistatic (evolving only on timescales ~ 1). Both approaches are equivalent to assuming that the exact metric contains no high-frequency oscillations occurring on the body's natural timescale $\sim \varepsilon$. In other words, the body is assumed to be in equilibrium. If we did not make this assumption, high-frequency oscillations could propagate throughout the external spacetime, invalidating our external expansion.

Since we are interested in the inner expansion only insofar as it informs the outer expansion, we shall not seek to explicitly solve the perturbative Einstein equation in the inner expansion. See Ref. [17] for the forms of the equations and an example of an explicit solution in the case of a perturbed black hole.

22 General expansion in the buffer region

We now seek the general solution to the equations of the outer expansion in the buffer region. To perform the expansion, we adopt Fermi coordinates centered about γ and expand for small s. In solving the first-order equations, we will determine $a^{(0)\mu}$; in solving the second-order equations, we will determine $a^{(1)\mu}$, including the self-force on the body. Although we perform this calculation in the Lorenz gauge, the choice of gauge is not essential for our purposes here — the essential aspect is our assumed expansion of the acceleration of the world line γ.

22.1 Metric expansions

The method of matched asymptotic expansions relies on the fact that the inner and outer expansion agree term by term when re-expanded in the buffer region, where $\varepsilon \ll s \ll 1$. To illustrate this idea of matching, consider the forms of the two expansions in the buffer region. The inner expansion holds $\tilde{s}$ constant (since $R \sim s$) while expanding for small ε. But if $\tilde{s}$ is replaced with its value s/ε, the inner expansion takes the form $\mathsf{g}_{\alpha\beta} = g^{\rm body}_{\alpha\beta}(s/\varepsilon)+\varepsilon H^{(1)}_{\alpha\beta}(s/\varepsilon)+\cdots$, where each term has a dependence on ε that can be expanded in the limit $\varepsilon \to 0$ to arrive at the schematic forms $g^{\rm body}(s/\varepsilon) = 1\oplus\varepsilon/s\oplus\varepsilon^2/s^2\oplus\ldots$ and $\varepsilon H^{(1)}(s/\varepsilon) = s\oplus\varepsilon\oplus\varepsilon^2/s\oplus\cdots$, where $\oplus$ signifies "plus terms of the form" and the expanded quantities can be taken to be components in Fermi coordinates. Here we have preemptively restricted the form of the expansions, since terms such as s^2/ε must vanish because they would have no corresponding terms in the outer expansion. Putting these two expansions together, we arrive at

$$\mathsf{g}(\text{buffer}) = 1 \oplus \frac{\varepsilon}{s} \oplus s \oplus \varepsilon \oplus \cdots . \tag{22.1}$$

Since this expansion relies on both an expansion at fixed $\tilde{s}$ and an expansion at fixed s, it can be expected to be accurate if $s \ll 1$ and $\varepsilon \ll s$ — that is, in the buffer region $\varepsilon \ll s \ll 1$.

On the other hand, the outer expansion holds s constant (since s is formally of the order of the global external coordinates) while expanding for small ε, leading to the form $\mathsf{g}_{\alpha\beta} = g_{\alpha\beta}(s)+\varepsilon h^{(1)}_{\alpha\beta}(s)+\cdots$. But very near the world line, each term in this expansion can be expanded for small s, leading to $g = 1\oplus s\oplus s^2\oplus\cdots$ and $\varepsilon h^{(1)} = \varepsilon/s\oplus\varepsilon\oplus\varepsilon s\oplus\cdots$. (Again, we have restricted this form because terms such as ε/s^2 cannot arise in the inner expansion.) Putting these two expansions together, we arrive at

$$\mathsf{g}(\text{buffer}) = 1 \oplus s \oplus \frac{\varepsilon}{s} \oplus \varepsilon \cdots . \tag{22.2}$$

Since this expansion relies on both an expansion at fixed s and an expansion for small s, it can be expected to be accurate in the buffer region $\varepsilon \ll s \ll 1$. As we can see, the two buffer-region expansions have an identical form; and because they are expansions of the same exact metric g, they must agree term by term.

One can make use of this fact by first determining the inner and outer expansions as fully as possible, then fixing any unknown functions in them by matching them term by term in the buffer region; this was the route taken in, e.g., Refs. [6, 15, 33, 170]. However, such an approach is complicated by the subtleties of matching in a diffeomorphism-invariant theory, where the inner and outer expansions are generically in different coordinate systems. See Ref. [18] for an analysis of the limitations of this approach as it has typically been implemented. Alternatively, one can take the opposite approach, working in the buffer region first, constraining the forms of the two expansions by making use of their matching, then using the buffer-region information to construct a global solution; this was the route taken in, e.g., Refs. [16, 17, 166]. In general, some mixture of these two approaches can be taken. Our calculation follows Ref. [17]. The only information we take from the inner expansion is its general form, which is characterized by the multipole moments of the body. From this information, we determine the external expansion, and thence the equation of motion of the world line.

Over the course of our calculation, we will find that the external metric perturbation in the buffer region is expressed as the sum of two solutions: one that formally diverges at $s = 0$ and is entirely determined from a combination of (i) the multipole moments of the internal background metric $g^{\rm body}_{\alpha\beta}$, (ii) the Riemann tensor of the external background $g_{\alpha\beta}$, and (iii) the acceleration of the world line γ; and a second solution that is formally regular at $s = 0$ and depends on the past history of the body and the initial conditions of

the field. At leading order, these two solutions are identified as the Detweiler-Whiting singular and regular fields $h^{\mathrm{S}}_{\alpha\beta}$ and $h^{\mathrm{R}}_{\alpha\beta}$, respectively, and the self-force is determined entirely by $h^{\mathrm{R}}_{\alpha\beta}$. Along with the self-force, the acceleration of the world line includes the Papapetrou spin-force [179]. This calculation leaves us with the self-force in terms of the the metric perturbation in the neighbourhood of the body. In Sec. 23, we use the local information from the buffer region to construct a global solution for the metric perturbation, completing the solution of the problem.

22.2 The form of the expansion

Before proceeding, we define some notation. We use the multi-index notation $\omega^L := \omega^{i_1}\cdots\omega^{i_\ell} := \omega^{i_1\cdots i_\ell}$. Angular brackets denote the STF combination of the enclosed indices, and a tensor bearing a hat is an STF tensor. To accomodate this, we now write the Fermi spatial coordinates as x^a, instead of $\hat{x}^a$ as they were written in previous sections. Finally, we define the one-forms $t_\alpha := \partial_\alpha t$ and $x^a_\alpha := \partial_\alpha x^a$.

One should note that the coordinate transformation $x^\alpha(t, x^a)$ between Fermi coordinates and the global coordinates is ε-dependent, since Fermi coordinates are tethered to an ε-dependent world line. If one were using a regular expansion, then this coordinate transformation would devolve into a background coordinate transformation to a Fermi coordinate system centered on a geodesic world line, combined with a gauge transformation to account for the ε-dependence. But in the self-consistent expansion, the transformation is purely a background transformation, because the ε-dependence in it is reducible to that of the fixed world line.

Because the dependence on ε in the coordinate transformation cannot be reduced to a gauge transformation, in Fermi coordinates the components $g_{\alpha\beta}$ of the background metric become ε-dependent. This dependence takes the explicit form of factors of the acceleration $a^\mu(t,\varepsilon)$ and its derivatives, for which we have assumed the expansion $a^i(t,\varepsilon) = a^{(0)i}(t) + a^{(1)i}(t;\gamma) + O(\varepsilon^2)$. There is also an implicit dependence on ε in that the proper time t on the world line depends on ε if written as a function of the global coordinates; but this dependence can be ignored so long as we work consistently with Fermi coordinates.

Of course, even in these ε-dependent coordinates, $g_{\mu\nu}$ remains the background metric of the outer expansion, and $h^{(n)}_{\mu\nu}$ is an exact solution to the wave equation (21.7). At first order we will therefore obtain $h^{(1)}_{\mu\nu}$ exactly in Fermi coordinates, for arbitrary a^μ. However, for some purposes an approximate solution of the wave equation may suffice, in which case we may utilize the expansion of a^μ. Substituting that expansion into $g_{\mu\nu}$ and $h^{(n)}_{\mu\nu}$ yields the *buffer-region expansions*

$$g_{\mu\nu}(t,x^a;a^i) = g_{\mu\nu}(t,x^a;a^{(0)i}) + \varepsilon g^{(1)}_{\mu\nu}(t,x^a;a^{(1)i}) + O(\varepsilon^2) \tag{22.3}$$

$$h^{(n)}_{\mu\nu}(t,x^a;a^i) = h^{(n)}_{\mu\nu}(t,x^a;a^{(0)i}) + O(\varepsilon), \tag{22.4}$$

where indices refer to Fermi coordinates, $g^{(1)}_{\mu\nu}$ is linear in $a^{(1)i}$ and its derivatives, and for future compactness of notation we define $h^{(n)}_{B\mu\nu}(t,x^a) := h^{(n)}_{\mu\nu}(t,x^a;a^{(0)i})$, where the subscript 'B' stands for 'buffer'. In the case that $a^{(0)i} = 0$, these expansions will significantly reduce the complexity of calculations in the buffer region. For that reason, we shall use them in solving the second-order wave equation, but we stress that they are simply a means of economizing calculations in Fermi coordinates; they do not play a fundamental role in the formalism, and one could readily do without them.

Now, we merely assume that in the buffer region there exists a smooth coordinate transformation between the local coordinates (T, R, Θ^A) and the Fermi coordinates (t, x^a) such that $T \sim t$, $R \sim s$, and $\Theta^A \sim \theta^A$. The buffer region corresponds to asymptotic infinity $s \gg \varepsilon$ (or $\tilde{s} \gg 1$) in the internal spacetime. So after re-expressing $\tilde{s}$ as s/ε, the internal background metric can be expanded as

$$g^{\mathrm{body}}_{\alpha\beta}(t,\tilde{s},\theta^A) = \sum_{n\geq 0}\left(\frac{\varepsilon}{s}\right)^n g^{\mathrm{body}(n)}_{\alpha\beta}(t,\theta^A). \tag{22.5}$$

As mentioned above, since the outer expansion has no negative powers of ε, we exclude them from the inner expansion. Furthermore, since $g_{\alpha\beta} + h_{\alpha\beta} = g^{\mathrm{body}}_{\alpha\beta} + H_{\alpha\beta}$, we must have $g^{\mathrm{body}(0)}_{\alpha\beta} = g_{\alpha\beta}(x^a = 0)$, since these are the only terms independent of both ε and s. Thus, noting that $g_{\alpha\beta}(x^a = 0) = \eta_{\alpha\beta} := \mathrm{diag}(-1,1,1,1)$,

we can write

$$g^{\rm body}_{\alpha\beta}(t,\tilde{s},\theta^A)=\eta_{\alpha\beta}+\frac{\varepsilon}{s}g^{\rm body(1)}_{\alpha\beta}(t,\theta^A)+\left(\frac{\varepsilon}{s}\right)^2 g^{\rm body(2)}_{B\alpha\beta}(t,\theta^A)+O(\varepsilon^3/s^3), \tag{22.6}$$

implying that the internal background spacetime is asymptotically flat.

We assume that the perturbation $H_{\alpha\beta}$ can be similarly expanded in powers of ε at fixed $\tilde{s}$,

$$H_{\alpha\beta}(t,\tilde{s},\theta^A,\varepsilon)=\varepsilon H^{(1)}_{\alpha\beta}(t,\tilde{s},\theta^A;\gamma)+\varepsilon^2 H^{(2)}_{\alpha\beta}(t,\tilde{s},\theta^A;\gamma)+O(\varepsilon^3), \tag{22.7}$$

and that each coefficient can be expanded in powers of $1/\tilde{s}=\varepsilon/s$ to yield

$$\varepsilon H^{(1)}_{\alpha\beta}(\tilde{s})=sH^{(0,1)}_{\alpha\beta}+\varepsilon H^{(1,0)}_{\alpha\beta}+\frac{\varepsilon^2}{s}H^{(2,-1)}_{\alpha\beta}+O(\varepsilon^3/s^2), \tag{22.8}$$

$$\varepsilon^2 H^{(2)}_{\alpha\beta}(\tilde{s})=s^2H^{(0,2)}_{\alpha\beta}+\varepsilon sH^{(1,1)}_{\alpha\beta}+\varepsilon^2 H^{(2,0)}_{\alpha\beta}+\varepsilon^2\ln s\, H^{(2,0,\ln)}_{\alpha\beta}+O(\varepsilon^3/s), \tag{22.9}$$

$$\varepsilon^3 H^{(3)}_{\alpha\beta}(\tilde{s})=O(\varepsilon^3,\varepsilon^2 s,\varepsilon s^2,s^3), \tag{22.10}$$

where $H^{(n,m)}_{\alpha\beta}$, the coefficient of ε^n and s^m, is a function of t and θ^A (and potentially a functional of γ). Again, the form of this expansion is constrained by the fact that no negative powers of ε can appear in the buffer region. (One might think that terms with negative powers of ε could be allowed in the expansion of $g^{\rm body}_{\alpha\beta}$ if they are exactly canceled by terms in the expansion of $H_{\alpha\beta}$, but the differing powers of s in the two expansions makes this impossible.) Note that explicit powers of s appear because $\varepsilon\tilde{s}=s$. Also note that we allow for a logarithmic term at second order in ε; this term arises because the retarded time in the internal background includes a logarithmic correction of the form $\varepsilon\ln s$ (e.g., $t-r\to t-r^*$ in Schwarzschild coordinates). Since we seek solutions to a wave equation, this correction to the characteristic curves induces a corresponding correction to the first-order perturbations.

The expansion of $H_{\alpha\beta}$ may or may not hold the acceleration fixed. Regardless of this choice, the general form of the expansion remains valid: incorporating the expansion of the acceleration would merely shuffle terms from one coefficient to another. And since the internal metric $g^{\rm body}_{\alpha\beta}+H_{\alpha\beta}$ must equal the external metric $g_{\alpha\beta}+h_{\alpha\beta}$, the general form of the above expansions of $g^{\rm body}_{\alpha\beta}$ and $H_{\alpha\beta}$ completely determines the general form of the external perturbations:

$$h^{(1)}_{\alpha\beta}=\frac{1}{s}h^{(1,-1)}_{\alpha\beta}+h^{(1,0)}_{\alpha\beta}+sh^{(1,1)}_{\alpha\beta}+O(s^2), \tag{22.11}$$

$$h^{(2)}_{\alpha\beta}=\frac{1}{s^2}h^{(2,-2)}_{\alpha\beta}+\frac{1}{s}h^{(2,-1)}_{\alpha\beta}+h^{(2,0)}_{\alpha\beta}+\ln s\, h^{(2,0,\ln)}_{\alpha\beta}+O(s), \tag{22.12}$$

where each $h^{(n,m)}_{\alpha\beta}$ depends only on t and θ^A, along with an implicit functional dependence on γ. If the internal expansion is performed with a^μ held fixed, then the internal and external quantities are related order by order: e.g., $\sum_m H^{(0,m)}_{\alpha\beta}=g_{\alpha\beta}$, $h^{(1,-1)}_{\alpha\beta}=g^{\rm body(1)}_{\alpha\beta}$, and $h^{(1,0)}_{\alpha\beta}=H^{(1,0)}_{\alpha\beta}$. Since we are not concerned with determining the internal perturbations, the only such relationship of interest is $h^{(n,\text{-}n)}_{\alpha\beta}=g^{\rm body(n)}_{\alpha\beta}$. This equality tells us that the most divergent, s^{-n} piece of the nth-order perturbation $h^{(n)}_{\alpha\beta}$ is defined entirely by the nth-order piece of the internal background metric $g^{\rm body}_{\alpha\beta}$, which is the metric of the body if it were isolated.

To obtain a general solution to the Einstein equation, we write each $h^{(n,m)}_{\alpha\beta}$ as an expansion in terms of irreducible symmetric trace-free pieces:

$$h^{(n,m)}_{tt}=\sum_{\ell\geq 0}\hat{A}^{(n,m)}_L\hat{\omega}^L, \tag{22.13}$$

$$h^{(n,m)}_{ta}=\sum_{\ell\geq 0}\hat{B}^{(n,m)}_L\hat{\omega}_a{}^L+\sum_{\ell\geq 1}\left[\hat{C}^{(n,m)}_{aL-1}\hat{\omega}^{L-1}+\epsilon_{ab}{}^c\hat{D}^{(n,m)}_{cL-1}\hat{\omega}^{bL-1}\right], \tag{22.14}$$

$$\begin{aligned}h^{(n,m)}_{ab}&=\delta_{ab}\sum_{\ell\geq 0}\hat{K}^{(n,m)}_L\hat{\omega}^L+\sum_{\ell\geq 0}\hat{E}^{(n,m)}_L\hat{\omega}_{ab}{}^L+\sum_{\ell\geq 1}\left[\hat{F}^{(n,m)}_{L-1\langle a}\hat{\omega}_{b\rangle}{}^{L-1}+\epsilon^{cd}{}_{(a}\hat{\omega}_{b)c}{}^{L-1}\hat{G}^{(n,m)}_{dL-1}\right]\\&\quad+\sum_{\ell\geq 2}\left[\hat{H}^{(n,m)}_{abL-2}\hat{\omega}^{L-2}+\epsilon^{cd}{}_{(a}\hat{I}^{(n,m)}_{b)dL-2}\hat{\omega}_c{}^{L-2}\right].\end{aligned} \tag{22.15}$$

Here a hat indicates that a tensor is STF with respect to δ_{ab}, angular brackets $\langle\rangle$ indicate the STF combination of enclosed indices, parentheses indicate the symmetric combination of enclosed indices, and symbols such as $\hat{A}_L^{(n,m)}$ are functions of time (and potentially functionals of γ) and are STF in all their indices. Each term in this expansion is linearly independent of all the other terms. All the quantities on the right-hand side are flat-space Cartesian tensors; their indices can be raised or lowered with δ_{ab}. Refer to Appendix B for more details about this expansion.

Now, since the wave equations (21.4) and (21.5) are covariant, they must still hold in the new coordinate system, despite the additional ε-dependence. Thus, both equations could be solved for arbitrary acceleration in the buffer region. However, due to the length of the calculations involved, we will instead solve the equations

$$E_{\alpha\beta}[h^{(1)}] = 0, \tag{22.16}$$

$$E^{(0)}_{\alpha\beta}[h^{(2)}_B] = 2\delta^2 R^{(0)}_{\alpha\beta}[h^{(1)}] + O(\varepsilon), \tag{22.17}$$

where $E^{(0)}[f] := E[f]\big|_{a=a_{(0)}}$ and $\delta^2 R^{(0)}[f] := \delta^2 R[f]\big|_{a=a_{(0)}}$. In analogy with the notation used for $L^{(n)}_\mu$, $E^{(1)}_{\mu\nu}[f]$ and $\delta^2 R^{(1)}_{\mu\nu}[f]$ would be linear in $a^\mu_{(1)}$, $E^{(2)}_{\mu\nu}[f]$ and $\delta^2 R^{(2)}_{\mu\nu}[f]$ would be linear in $a^\mu_{(2)}$ and quadratic in $a^\mu_{(1)}$, and so on. For a function $f \sim 1$, $L^{(n)}_\mu[f]$, $E^{(n)}_{\mu\nu}[f]$, and $\delta^2 R^{(n)}_{\mu\nu}[f]$ correspond to the coefficients of ε^n in expansions in powers of ε. Equation (22.16) is identical to Eq. (21.4). Equation (22.17) follows directly from substituting Eqs. (22.3) and (22.4) into Eq. (21.5); in the buffer region, it captures the dominant behaviour of $h^{(2)}_{\alpha\beta}$, represented by the approximation $h^{(2)}_{B\alpha\beta}$, but it does not capture its full dependence on acceleration. If one desired a global second-order solution, one might need to solve Eq. (21.5), but for our purpose, which is to determine the first-order acceleration $a^\mu_{(1)}$, Eq. (22.17) will suffice.

Unlike the wave equations, the gauge conditions (21.10) and (21.11) already incorporate the expansion of the acceleration. As such, they are unmodified by the replacement of the second-order wave equation (21.5) with its approximation (22.17). So we can write

$$L^{(0)}_\mu\big[h^{(1)}\big] = 0, \tag{22.18}$$

$$L^{(1)}_\mu\big[h^{(1)}\big] = -L^{(0)}_\mu\big[h^{(2)}_B\big], \tag{22.19}$$

where the first equation is identical to Eq. (21.10), and the second to Eq. (21.11). (The second identity holds because $L^{(0)}_\mu\big[h^{(2)}_B\big] = L^{(0)}_\mu\big[h^{(2)}\big]$, since $h^{(2)}_{B\alpha\beta}$ differs from $h^{(2)}_{\alpha\beta}$ by $a^\alpha_{(1)}$ and higher acceleration terms, which are set to zero in $L^{(0)}_\mu$.) We remind the reader that while this gauge choice is important for finding the external perturbations globally, any other choice would suffice in the buffer region calculation.

In what follows, the reader may safely assume that all calculations are lengthy unless noted otherwise.

22.3 First-order solution in the buffer region

In principle, solving the first-order Einstein equation in the buffer region is straightforward. One need simply substitute the expansion of $h^{(1)}_{\alpha\beta}$, given in Eq. (22.11), into the linearized wave equation (22.16) and the gauge condition (22.18). Equating powers of s in the resulting expansions then yields a sequence of equations that can be solved for successively higher-order terms in $h^{(1)}_{\alpha\beta}$. Solving these equations consists primarily of expressing each quantity in its irreducible STF form, using the decompositions (B.3) and (B.7); since the terms in this STF decomposition are linearly independent, we can solve each equation term by term. This calculation is aided by the fact that $\nabla_\alpha = x^a_\alpha \partial_a + O(s^0)$, so that, for example, the wave operator $E_{\alpha\beta}$ consists of a flat-space Laplacian $\partial^a\partial_a$ plus corrections of order $1/s$. Appendix B also lists many useful identities, particularly $\partial^a s = \omega^a := x^a/s$, $\omega^a\partial_a\hat{\omega}^L = 0$, and the fact that $\hat{\omega}^L$ is an eigenvector of the flat-space Laplacian: $s^2\partial^a\partial_a\hat{\omega}^L = -\ell(\ell+1)\hat{\omega}^L$.

Summary of results

Before proceeding with the calculation, which consists mostly of tedious and lengthy algebra, we summarize the results. The first-order perturbation $h^{(1)}_{\alpha\beta}$ consists of two pieces, which we will eventually identify with

the Detweiler-Whiting regular and singular fields. In the buffer-region expansion, the regular field consists entirely of unknowns, which is to be expected since as a free radiation field, it must be provided by boundary data. Only when we consider the global solution, in Sec. 23, will we express it in terms of a tail integral. On the other hand, the singular field is locally determined, and it is characterized by the body's monopole moment m. More precisely, it is fully determined by the tidal fields of the external background spacetime and the Arnowitt-Deser-Misner mass of the internal background spacetime $g^{\rm body}_{\alpha\beta}$. By itself the wave equation does not restrict the behaviour of this monopole moment, but imposing the gauge condition produces the evolution equations

$$\partial_t m = 0, \qquad a^i_{(0)} = 0. \tag{22.20}$$

Hence, at leading order, the body behaves as a test particle, with constant mass and vanishing acceleration.

Order $(1,-1)$

We now proceed to the details of the calculation. We begin with the most divergent term in the wave equation: the order $1/s^3$, flat-space Laplacian term

$$\frac{1}{s}\partial^c\partial_c h^{(1,-1)}_{\alpha\beta} = 0. \tag{22.21}$$

The tt-component of this equation is

$$0 = -\sum_{\ell\geq 0}\ell(\ell+1)\hat{A}^{(1,-1)}_L\hat{\omega}^L, \tag{22.22}$$

from which we read off that $\hat{A}^{(1,-1)}$ is arbitrary and $\hat{A}^{(1,-1)}_L$ must vanish for all $\ell\geq 1$. The ta-component is

$$\begin{aligned}0 = &-\sum_{\ell\geq 0}(\ell+1)(\ell+2)\hat{B}^{(1,-1)}_L\hat{\omega}_a{}^L - \sum_{\ell\geq 1}\ell(\ell-1)\hat{C}^{(1,-1)}_{aL-1}\hat{\omega}^{L-1}\\ &-\sum_{\ell\geq 1}\ell(\ell+1)\epsilon_{abc}\hat{D}^{(1,-1)}_{cL-1}\hat{\omega}_b{}^{L-1},\end{aligned} \tag{22.23}$$

from which we read off that $\hat{C}^{(1,-1)}_a$ is arbitrary and all other coefficients must vanish. Lastly, the ab-component is

$$\begin{aligned}0 = &-\delta_{ab}\sum_{\ell\geq 0}\ell(\ell+1)\hat{K}^{(1,-1)}_L\hat{\omega}^L - \sum_{\ell\geq 0}(\ell+2)(\ell+3)\hat{E}^{(1,-1)}_L\hat{\omega}_{ab}{}^L\\ &-\sum_{\ell\geq 1}\ell(\ell+1)\hat{F}^{(1,-1)}_{L-1\langle a}\hat{\omega}_{b\rangle}{}^{L-1} - \sum_{\ell\geq 1}(\ell+1)(\ell+2)\epsilon_{cd(a}\hat{\omega}_{b)}{}^{cL-1}\hat{G}^{(1,-1)}_{dL-1}\\ &-\sum_{\ell\geq 2}(\ell-2)(\ell-1)\hat{H}^{(1,-1)}_{abL-2}\hat{\omega}^{L-2} - \sum_{\ell\geq 2}\ell(\ell-1)\epsilon_{cd(a}\hat{I}^{(1,-1)}_{b)dL-2}\hat{\omega}_c{}^{L-2},\end{aligned} \tag{22.24}$$

from which we read off that $\hat{K}^{(1,-1)}$ and $\hat{H}^{(1,-1)}_{ab}$ are arbitrary and all other coefficients must vanish. Thus, we find that the wave equation constrains $h^{(1,-1)}_{\alpha\beta}$ to be

$$h^{(1,-1)}_{\alpha\beta} = \hat{A}^{(1,-1)}t_\alpha t_\beta + 2\hat{C}^{(1,-1)}_a t_{(\beta}x^a_{\alpha)} + (\delta_{ab}\hat{K}^{(1,-1)} + \hat{H}^{(1,-1)}_{ab})x^a_\alpha x^b_\beta. \tag{22.25}$$

This is further constrained by the most divergent, $1/s^2$ term in the gauge condition, which reads

$$-\frac{1}{s^2}h^{(1,-1)}_{\alpha c}\omega^c + \frac{1}{2s^2}\omega_\alpha\eta^{\mu\nu}h^{(1,-1)}_{\mu\nu} = 0. \tag{22.26}$$

From the t-component of this equation, we read off $\hat{C}^{(1,-1)}_a = 0$; from the a-component, $\hat{K}^{(1,-1)} = \hat{A}^{(1,-1)}$ and $\hat{H}^{(1,-1)}_{ab} = 0$. Thus, $h^{(1,-1)}_{\alpha\beta}$ depends only on a single function of time, $\hat{A}^{(1,-1)}$. By the definition of the

ADM mass, this function (times ε) must be twice the mass of the internal background spacetime. Thus, $h^{(1,-1)}_{\alpha\beta}$ is fully determined to be

$$h^{(1,-1)}_{\alpha\beta} = 2m(t)(t_\alpha t_\beta + \delta_{ab}x^a_\alpha x^b_\beta), \tag{22.27}$$

where $m(t)$ is defined to be the mass at time t divided by the initial mass $\varepsilon := m_0$. (Because the mass will be found to be a constant, $m(t)$ is merely a placeholder; it is identically unity. We could instead set ε equal to unity at the end of the calculation, in which case m would simply be the mass at time t. Obviously, the difference between the two approaches is immaterial.)

Order $(1,0)$

At the next order, $h^{(1,0)}_{\alpha\beta}$, along with the acceleration of the world line and the time-derivative of the mass, first appears in the Einstein equation. The order $1/s^2$ term in the wave equation is

$$\partial^c\partial_c h^{(1,0)}_{\alpha\beta} = -\frac{2m}{s^2}a_c\omega^c(3t_\alpha t_\beta - \delta_{ab}x^a_\alpha x^b_\beta), \tag{22.28}$$

where the terms on the right arise from the wave operator acting on $s^{-1}h^{(1,-1)}_{\alpha\beta}$. This equation constrains $h^{(1,0)}_{\alpha\beta}$ to be

$$\begin{aligned} h^{(1,0)}_{tt} &= \hat{A}^{(1,0)} + 3ma_c\omega^c, \\ h^{(1,0)}_{ta} &= \hat{C}^{(1,0)}_a, \\ h^{(1,0)}_{ab} &= \delta_{ab}\left(\hat{K}^{(1,0)} - ma_c\omega^c\right) + \hat{H}^{(1,0)}_{ab}. \end{aligned} \tag{22.29}$$

Substituting this result into the order $1/s$ term in the gauge condition, we find

$$-\frac{4}{s}t_\alpha\partial_t m + \frac{4m}{s}a^{(0)}_a x^a_\alpha = 0. \tag{22.30}$$

Thus, both the leading-order part of the acceleration and the rate of change of the mass of the body vanish:

$$\frac{\partial m}{\partial t} = 0\,, \qquad a^i_{(0)} = 0. \tag{22.31}$$

Order $(1,1)$

At the next order, $h^{(1,1)}_{\alpha\beta}$, along with squares and derivatives of the acceleration, first appears in the Einstein equation, and the tidal fields of the external background couple to $s^{-1}h^{(1,-1)}_{\alpha\beta}$. The order $1/s$ term in the wave equation becomes

$$\left(s\partial^c\partial_c + \frac{2}{s}\right)h^{(1,1)}_{tt} = -\frac{20m}{3s}\mathcal{E}_{ij}\hat{\omega}^{ij} - \frac{3m}{s}a_{\langle i}a_{j\rangle}\hat{\omega}^{ij} + \frac{8m}{s}a_ia^i, \tag{22.32}$$

$$\left(s\partial^c\partial_c + \frac{2}{s}\right)h^{(1,1)}_{ta} = -\frac{8m}{3s}\epsilon_{aij}\mathcal{B}^j_k\hat{\omega}^{ik} - \frac{4m}{s}\dot{a}_a, \tag{22.33}$$

$$\begin{aligned} \left(s\partial^c\partial_c + \frac{2}{s}\right)h^{(1,1)}_{ab} = {} & \frac{20m}{9s}\delta_{ab}\mathcal{E}_{ij}\hat{\omega}^{ij} - \frac{76m}{9s}\mathcal{E}_{ab} - \frac{16m}{3s}\mathcal{E}^i_{\langle a}\hat{\omega}_{b\rangle i} + \frac{8m}{s}a_{\langle a}a_{b\rangle} \\ & + \frac{m}{s}\delta_{ab}\left(\tfrac{8}{3}a_ia^i - 3a_{\langle i}a_{j\rangle}\hat{\omega}^{ij}\right). \end{aligned} \tag{22.34}$$

From the tt-component, we read off that $\hat{A}^{(1,1)}_i$ is arbitrary, $\hat{A}^{(1,1)} = 4ma_ia^i$, and $\hat{A}^{(1,1)}_{ij} = \frac{5}{3}m\mathcal{E}_{ij} + \frac{3}{4}ma_{\langle i}a_{j\rangle}$; from the ta-component, $\hat{B}^{(1,1)}$, $\hat{C}^{(1,1)}_{ij}$, and $\hat{D}^{(1,1)}_i$ are arbitrary, $\hat{C}^{(1,1)}_i = -2m\dot{a}_i$, and $\hat{D}^{(1,1)}_{ij} = \frac{2}{3}m\mathcal{B}_{ij}$; from the ab component, $\hat{K}^{(1,1)}_i$, $\hat{F}^{(1,1)}_i$, $\hat{H}^{(1,1)}_{ijk}$, and $\hat{I}^{(1,1)}_{ij}$ are arbitrary, and $\hat{K}^{(1,1)} = \frac{4}{3}ma_ia^i$, $\hat{K}^{(1,1)}_{ij} = -\frac{5}{9}m\mathcal{E}_{ij} + \frac{3}{4}ma_{\langle i}a_{j\rangle}$, $\hat{F}^{(1,1)}_{ij} = \frac{4}{3}m\mathcal{E}_{ij}$, and $\hat{H}^{(1,1)}_{ij} = -\frac{38}{9}m\mathcal{E}_{ij} + 4ma_{\langle i}a_{j\rangle}$.

Substituting this into the order s^0 terms in the gauge condition, we find

$$\begin{aligned}0 &= (\omega^i + s\partial^i)h^{(1,1)}_{\alpha i} - \tfrac{1}{2}\eta^{\mu\nu}(\omega_a - s\partial_a)h^{(1,1)}_{\mu\nu}x^a_\alpha - \partial_t h^{(1,0)}_{\alpha t} - \tfrac{1}{2}\eta^{\mu\nu}\partial_t h^{(1,0)}_{\mu\nu}t_\alpha \\ &\quad + \tfrac{4}{3}m\mathcal{E}_{ij}\hat{\omega}^{ij}\omega_\alpha + \tfrac{2}{3}m\mathcal{E}_{ai}\omega^i x^a_\alpha,\end{aligned} \tag{22.35}$$

where the equation is to be evaluated at $a^i = a^i_{(0)} = 0$. From the t-component, we read off

$$\hat{B}^{(1,1)} = \tfrac{1}{6}\partial_t\left(\hat{A}^{(1,0)} + 3\hat{K}^{(1,0)}\right). \tag{22.36}$$

From the a-component,

$$\hat{F}^{(1,1)}_a = \tfrac{3}{10}\left(\hat{K}^{(1,1)}_a - \hat{A}^{(1,1)}_a + \partial_t\hat{C}^{(1,0)}_a\right). \tag{22.37}$$

It is understood that both these equations hold only when evaluated at $a^i = 0$.

Thus, the order s component of $h^{(1)}_{\alpha\beta}$ is

$$\begin{aligned}h^{(1,1)}_{tt} &= 4ma_i a^i + \hat{A}^{(1,1)}_i\omega^i + \tfrac{5}{3}m\mathcal{E}_{ij}\hat{\omega}^{ij} + \tfrac{3}{4}ma_{\langle i}a_{j\rangle}\hat{\omega}^{ij}, \\ h^{(1,1)}_{ta} &= \hat{B}^{(1,1)}\omega_a - 2m\dot{a}_a + \hat{C}^{(1,1)}_{ai}\omega^i + \epsilon_{ai}{}^j\hat{D}^{(1,1)}_j\omega^i + \tfrac{2}{3}m\epsilon_{aij}\mathcal{B}^j_k\hat{\omega}^{ik}, \\ h^{(1,1)}_{ab} &= \delta_{ab}\big(\tfrac{4}{3}ma_i a^i + \hat{K}^{(1,1)}_i\omega^i - \tfrac{5}{9}m\mathcal{E}_{ij}\hat{\omega}^{ij} + \tfrac{3}{4}ma_{\langle i}a_{j\rangle}\hat{\omega}^{ij}\big) + \tfrac{4}{3}m\mathcal{E}^i_{\langle a}\hat{\omega}_{b\rangle i} \\ &\quad - \tfrac{38}{9}m\mathcal{E}_{ab} + 4ma_{\langle a}a_{b\rangle} + \hat{H}^{(1,1)}_{abi}\omega^i + \epsilon_i{}^j{}_{(a}\hat{I}^{(1,1)}_{b)j}\omega^i + \hat{F}^{(1,1)}_{\langle a}\omega_{b\rangle}.\end{aligned} \tag{22.38}$$

where $\hat{B}^{(1,1)}$ and $\hat{F}^{(1,1)}_a$ are constrained to satisfy Eqs. (22.36) and (22.37).

First-order solution

To summarize the results of this section, we have $h^{(1)}_{\alpha\beta} = s^{-1}h^{(1,-1)}_{\alpha\beta} + h^{(1,0)}_{\alpha\beta} + sh^{(1,1)}_{\alpha\beta} + O(s^2)$, where $h^{(1,-1)}_{\alpha\beta}$ is given in Eq. (22.27), $h^{(1,0)}_{\alpha\beta}$ is given in Eq. (22.29), and $h^{(1,1)}_{\alpha\beta}$ is given in Eq. (22.38). In addition, we have determined that the ADM mass of the internal background spacetime is time-independent, and that the acceleration of the body's world line vanishes at leading order.

22.4 Second-order solution in the buffer region

Though the calculations are much lengthier, solving the second-order Einstein equation in the buffer region is essentially no different from solving the first. We seek to solve the approximate wave equation (22.17), along with the gauge condition (22.19), for the second-order perturbation $h^{(2)}_{B\alpha\beta} := h^{(2)}_{\alpha\beta}\big|_{a=a_0}$; doing so will also, more importantly, determine the acceleration $a^\mu_{(1)}$. In this calculation, the acceleration is set to $a^i = a^i_{(0)} = 0$ everywhere except in the left-hand side of the gauge condition, $L^{(1)}_\mu[h^{(1)}]$, which is linear in $a^\mu_{(1)}$.

Summary of results

We first summarize the results. As at first order, the metric perturbation contains a regular, free radiation field and a singular, bound field; but in addition to these pieces, it also contains terms sourced by the first-order perturbation. Again, the regular field requires boundary data to be fully determined. And again, the singular field is characterized by the multipole moments of the body: the mass dipole M_i of the internal background metric $g^{\rm body}_{\alpha\beta}$, which measures the shift of the body's centre of mass relative to the world line; the spin dipole S_i of $g^{\rm body}_{\alpha\beta}$, which measures the spin of the body about the world line; and an effective correction δm to the body's mass. The wave equation by itself imposes no restriction on these quantities, but by imposing the gauge condition we find the evolution equations

$$\partial_t\delta m = \frac{m}{3}\partial_t\hat{A}^{(1,0)} + \frac{5m}{6}\partial_t\hat{K}^{(1,0)}, \tag{22.39}$$

$$\partial_t S_a = 0, \tag{22.40}$$

$$\partial^2_t M_a + \mathcal{E}_{ab}M^b = -a^{(1)}_a + \tfrac{1}{2}\hat{A}^{(1,1)}_a - \partial_t\hat{C}^{(1,0)}_a - \frac{1}{m}S_i\mathcal{B}^i_a. \tag{22.41}$$

The first of these tells us that the free radiation field created by the body creates a time-varying shift in the body's mass. We can immediately integrate it to find

$$\delta m(t) = \delta m(0) + \tfrac{1}{6}m\left[2\hat{A}^{(1,0)}(t) + 5\hat{K}^{(1,0)}(t)\right] - \tfrac{1}{6}m\left[2\hat{A}^{(1,0)}(0) + 5\hat{K}^{(1,0)}(0)\right]. \tag{22.42}$$

We note that this mass correction is entirely gauge dependent; it could be removed by redefining the time coordinate on the world line. In addition, one could choose to incorporate $\delta m(0)$ into the leading-order mass m. The second of the equations tells us that the body's spin is constant at this order; at higher orders, time-dependent corrections to the spin dipole would arise. The last of the equations is the principal result of this section. It tells us that the relationship between the acceleration of the world line and the drift of the body away from it is governed by (i) the local curvature of the background spacetime, as characterized by $\mathcal{E}_{ab}$ — this is the same term that appears in the geodesic deviation equation — (ii) the coupling of the body's spin to the local curvature — this is the Papapetrou spin force [179] — and (iii) the free radiation field created by the body — this is the self-force. We identify the world line as the body's by the condition $M_i = 0$. If we start with initial conditions $M_i(0) = 0 = \partial_t M_i(0)$, then the mass dipole remains zero for all times if and only if the world line satisfies the equation

$$a_a^{(1)} = \tfrac{1}{2}\hat{A}_a^{(1,1)} - \partial_t \hat{C}_a^{(1,0)} - \frac{1}{m}S_i\mathcal{B}^i_a. \tag{22.43}$$

This is the equation of motion we sought. It, along with the more general equation containing M_i, will be discussed further in the following section.

Order $(2,-2)$

We now proceed to the details of the calculation. Substituting the expansion

$$h_{\alpha\beta}^{(2)} = \frac{1}{s^2}h_{B\alpha\beta}^{(2,-2)} + \frac{1}{s}h_{B\alpha\beta}^{(2,-1)} + h_{B\alpha\beta}^{(2,0)} + \ln(s)h_{B\alpha\beta}^{(2,0,\ln)} + O(\varepsilon, s) \tag{22.44}$$

and the results for $h_{\alpha\beta}^{(1)}$ from the previous section into the wave equation and the gauge condition again yields a sequence of equations that can be solved for coefficients of successively higher-order powers (and logarithms) of s. Due to its length, the expansion of the second-order Ricci tensor is given in Appendix A. Note that since the approximate wave equation (22.17) contains an explicit $O(\varepsilon)$ correction, $h_{\alpha\beta}^{(2)}$ will be determined only up to $O(\varepsilon)$ corrections. For simplicity, we omit these $O(\varepsilon)$ symbols from the equations in this section; note, however, that these corrections do not effect the gauge condition, as discussed above.

To begin, the most divergent, order $1/s^4$ term in the wave equation reads

$$\frac{1}{s^4}\left(2 + s^2\partial^c\partial_c\right)h_{B\alpha\beta}^{(2,-2)} = \frac{4m^2}{s^4}\left(7\hat{\omega}_{ab} + \tfrac{4}{3}\delta_{ab}\right)x_\alpha^a x_\beta^b - \frac{4m^2}{s^4}t_\alpha t_\beta, \tag{22.45}$$

where the right-hand side is the most divergent part of the second-order Ricci tensor, as given in Eq. (A.3). From the tt-component of this equation, we read off $\hat{A}^{(2,-2)} = -2m^2$ and that $\hat{A}_a^{(2,-2)}$ is arbitrary. From the ta-component, $\hat{B}^{(2,-2)}$, $\hat{C}_{ab}^{(2,-2)}$, and $\hat{D}_c^{(2,-2)}$ are arbitrary. From the ab-component, $\hat{K}^{(2,-2)} = \frac{8}{3}m^2$, $\hat{E}^{(2,-2)} = -7m^2$, and $\hat{K}_a^{(2,-2)}$, $\hat{F}_a^{(2,-2)}$, $\hat{H}_{abc}^{(2,-2)}$, and $\hat{I}_{ab}^{(2,-2)}$ are arbitrary.

The most divergent, order $1/s^3$ terms in the gauge condition similarly involve only $h_{\alpha\beta}^{(2,-2)}$; they read

$$\frac{1}{s^3}\left(s\partial^b - 2\omega^b\right)h_{B\alpha b}^{(2,-2)} - \frac{1}{2s^3}\eta^{\mu\nu}x_\alpha^a\left(s\partial_a - 2\omega_a\right)h_{B\mu\nu}^{(2,-2)} = 0. \tag{22.46}$$

After substituting the results from the wave equation, the t-component of this equation determines that $\hat{C}_{ab}^{(2,-2)} = 0$. The a-component determines that $\hat{H}_{abc}^{(2,-2)} = 0$, $\hat{I}_{ab}^{(2,-2)} = 0$, and

$$\hat{F}_a^{(2,-2)} = 3\hat{K}_a^{(2,-2)} - 3\hat{A}_a^{(2,-2)}. \tag{22.47}$$

Thus, the order $1/s^2$ part of $h^{(2)}_{\alpha\beta}$ is given by

$$\begin{aligned} h^{(2,-2)}_{Btt} &= -2m^2 + \hat{A}^{(2,-2)}_i \omega^i, \\ h^{(2,-2)}_{Bta} &= \hat{B}^{(2,-2)}\omega_a + \epsilon_a{}^{ij}\omega_i \hat{D}^{(2,-2)}_j, \\ h^{(2,-2)}_{Bab} &= \delta_{ab}\left(\tfrac{8}{3}m^2 + \hat{K}^{(2,-2)}_i\omega^i\right) - 7m^2\hat{\omega}_{ab} + \hat{F}^{(2,-2)}_{\langle a}\omega_{b\rangle}, \end{aligned} \tag{22.48}$$

where $\hat{F}^{(2,-2)}_a$ is given by Eq. (22.47).

The metric perturbation in this form depends on five free functions of time. However, from calculations in flat spacetime, we know that order ε^2/s^2 terms in the metric perturbation can be written in terms of two free functions: a mass dipole and a spin dipole. We transform the perturbation into this "canonical" form by performing a gauge transformation (c.f. Ref. [180]). The transformation is generated by $\xi_\alpha = -\frac{1}{s}\hat{B}^{(2,-2)}t_\alpha - \frac{1}{2s}\hat{F}^{(2,-2)}_a x^a_\alpha$, the effect of which is to remove $\hat{B}^{(2,-2)}$ and $\hat{F}^{(2,-2)}_a$ from the metric. This transformation is a refinement of the Lorenz gauge. (Effects at higher order in ε and s will be automatically incorporated into the higher-order perturbations.) The condition $\hat{F}^{(2,-2)}_a - 3\hat{K}^{(2,-2)}_a + 3\hat{A}^{(2,-2)}_a = 0$ then becomes $\hat{K}^{(2,-2)}_a = \hat{A}^{(2,-2)}_a$. The remaining two functions are related to the ADM momenta of the internal spacetime:

$$\hat{A}^{(2,-2)}_i = 2M_i\,, \qquad \hat{D}^{(2,-2)}_i = 2S_i, \tag{22.49}$$

where M_i is such that $\partial_t M_i$ is proportional to the ADM linear momentum of the internal spacetime, and S_i is the ADM angular momentum. M_i is a mass dipole term; it is what would result from a transformation $x^a \to x^a + M^a/m$ applied to the $1/s$ term in $h^{(1)}_{\alpha\beta}$. S_i is a spin dipole term. Thus, the order $1/s^2$ part of $h^{(2)}_{B\alpha\beta}$ reads

$$\begin{aligned} h^{(2,-2)}_{Btt} &= -2m^2 + 2M_i\omega^i, \\ h^{(2,-2)}_{Bta} &= 2\epsilon_{aij}\omega^i S^j, \\ h^{(2,-2)}_{Bab} &= \delta_{ab}\left(\tfrac{8}{3}m^2 + 2M_i\omega^i\right) - 7m^2\hat{\omega}_{ab}. \end{aligned} \tag{22.50}$$

Order $(2,-1)$

At the next order, $1/s^3$, because the acceleration is set to zero, $h^{(2,-2)}_{B\alpha\beta}$ does not contribute to $E^{(0)}_{\mu\nu}[h^{(2)}]$, and $h^{(1,-1)}_{B\alpha\beta}$ does not contribute to $\delta^2 R^{(0)}_{\mu\nu}[h^{(1)}]$. The wave equation hence reads

$$\frac{1}{s}\partial^c\partial_c h^{(2,-1)}_{B\alpha\beta} = \frac{2}{s^3}\delta^2 R^{(0,-3)}_{\alpha\beta}\left[h^{(1)}\right], \tag{22.51}$$

where $\delta^2 R^{(0,-3)}_{\alpha\beta}\left[h^{(1)}\right]$ is given in Eqs. (A.4)–(A.6). The tt-component of this equation implies $s^2\partial^c\partial_c h^{(2,-1)}_{Btt} = 6m\hat{H}^{(1,0)}_{ij}\hat{\omega}^{ij}$, from which we read off that $\hat{A}^{(2,-1)}$ is arbitrary and $\hat{A}^{(2,-1)}_{ij} = -m\hat{H}^{(1,0)}_{ij}$. The ta-component implies $s^2\partial^c\partial_c h^{(2,-1)}_{Bta} = 6m\hat{C}^{(1,0)}_i\hat{\omega}^i_a$, from which we read off $\hat{B}^{(2,-1)}_i = -m\hat{C}^{(1,0)}_i$ and that $\hat{C}^{(2,-1)}_a$ is arbitrary. The ab-component implies

$$s^2\partial^c\partial_c h^{(2,-1)}_{Bab} = 6m\left(\hat{A}^{(1,0)} + \hat{K}^{(1,0)}\right)\hat{\omega}_{ab} - 12m\hat{H}^{(1,0)}_{i\langle a}\hat{\omega}_{b\rangle}{}^i + 2m\delta_{ab}\hat{H}^{(1,0)}_{ij}\hat{\omega}^{ij}, \tag{22.52}$$

from which we read off that $\hat{K}^{(2,-1)}$ is arbitrary, $\hat{K}^{(2,-1)}_{ij} = -\frac{1}{3}m\hat{H}^{(1,0)}_{ij}$, $\hat{E}^{(2,-1)} = -m\hat{A}^{(1,0)} - m\hat{K}^{(1,0)}$, $\hat{F}^{(2,-1)}_{ab} = 2m\hat{H}^{(1,0)}_{ab}$, and $\hat{H}^{(2,-1)}_{ab}$ is arbitrary. This restricts $h^{(2,-1)}_{\alpha\beta}$ to the form

$$\begin{aligned} h^{(2,-1)}_{Btt} &= \hat{A}^{(2,-1)} - m\hat{H}^{(1,0)}_{ij}\hat{\omega}^{ij}, \\ h^{(2,-1)}_{Bta} &= -m\hat{C}^{(1,0)}_i\hat{\omega}^i_a + \hat{C}^{(2,-1)}_a, \\ h^{(2,-1)}_{Bab} &= \delta_{ab}\left(\hat{K}^{(2,-1)} - \tfrac{1}{3}m\hat{H}^{(1,0)}_{ij}\hat{\omega}^{ij}\right) - m\left(\hat{A}^{(1,0)} + \hat{K}^{(1,0)}\right)\hat{\omega}_{ab} + 2m\hat{H}^{(1,0)}_{i\langle a}\hat{\omega}_{b\rangle}{}^i + \hat{H}^{(2,-1)}_{ab}. \end{aligned} \tag{22.53}$$

We next substitute $h^{(2,-2)}_{B\alpha\beta}$ and $h^{(2,-1)}_{B\alpha\beta}$ into the order $1/s^2$ terms in the gauge condition. The t-component becomes

$$\frac{1}{s^2}\left(4m\hat{C}^{(1,0)}_i + 12\partial_t M_i + 3\hat{C}^{(2,-1)}_i\right)\omega^i = 0, \tag{22.54}$$

from which we read off

$$\hat{C}^{(2,-1)}_i = -4\partial_t M_i - \tfrac{4}{3}m\hat{C}^{(1,0)}_i. \tag{22.55}$$

And the a-component becomes

$$\begin{aligned}0 = &\left(-\tfrac{4}{3}m\hat{A}^{(1,0)} - \tfrac{4}{3}m\hat{K}^{(1,0)} - \tfrac{1}{2}\hat{A}^{(2,-1)} + \tfrac{1}{2}\hat{K}^{(2,-1)}\right)\omega_a + \left(\tfrac{2}{3}m\hat{H}^{(1,0)}_{ai} - \hat{H}^{(2,-1)}_{ai}\right)\omega^i\\ &- 2\epsilon_{ija}\omega^i\partial_t S^j,\end{aligned} \tag{22.56}$$

from which we read off

$$\hat{A}^{(2,-1)} = \hat{K}^{(2,-1)} - \tfrac{8}{3}m\left(\hat{A}^{(1,0)} + \hat{K}^{(1,0)}\right), \tag{22.57}$$

$$\hat{H}^{(2,-1)}_{ij} = \tfrac{2}{3}m\hat{H}^{(1,0)}_{ij}, \tag{22.58}$$

and that the angular momentum of the internal background is constant at leading order:

$$\partial_t S^i = 0. \tag{22.59}$$

Thus, the order $1/s$ term in $h^{(2)}_{B\alpha\beta}$ is given by

$$\begin{aligned}h^{(2,-1)}_{Btt} &= \hat{K}^{(2,-1)} - \tfrac{8}{3}m\left(\hat{A}^{(1,0)} + \hat{K}^{(1,0)}\right) - m\hat{H}^{(1,0)}_{ij}\hat{\omega}^{ij},\\ h^{(2,-1)}_{Bta} &= -m\hat{C}^{(1,0)}_i\hat{\omega}^i_a - 4\partial_t M_i - \tfrac{4}{3}m\hat{C}^{(2,-1)}_i,\\ h^{(2,-1)}_{Bab} &= \delta_{ab}\left(\hat{K}^{(2,-1)} - \tfrac{1}{3}m\hat{H}^{(1,0)}_{ij}\hat{\omega}^{ij}\right) - m\left(\hat{A}^{(1,0)} + \hat{K}^{(1,0)}\right)\hat{\omega}_{ab} + 2m\hat{H}^{(1,0)}_{i\langle a}\hat{\omega}_{b\rangle}{}^i + \tfrac{2}{3}m\hat{H}^{(1,0)}_{ab}.\end{aligned} \tag{22.60}$$

Note that the undetermined function $\hat{K}^{(2,-1)}$ appears in precisely the form of a mass monopole. The value of this function will never be determined (though its time-dependence will be). This ambiguity arises because the mass m that we have defined is the mass of the internal *background* spacetime, which is based on the inner limit that holds ε/R fixed. A term of the form ε^2/R appears as a perturbation of this background, even when, as in this case, it is part of the mass monopole of the body. This is equivalent to the ambiguity in any expansion in one's choice of small parameter: one could expand in powers of ε, or one could expand in powers of $\varepsilon+\varepsilon^2$, and so on. It is also equivalent to the ambiguity in defining the mass of a non-isolated body; whether the "mass" of the body is taken to be m or $m+\frac{1}{2}\hat{K}^{(2,-1)}$ is a matter of taste. As we shall discover, the time-dependent part of $\hat{K}^{(2,-1)}$ is constructed from the tail terms in the first-order metric perturbation. Hence, the ambiguity in the definition of the mass is, at least in part, equivalent to whether or not one chooses to include the free gravitational field induced by the body in what one calls its mass. (In fact, any order ε incoming radiation, not just that originally produced by the body, will contribute to this effective mass.) In any case, we will define the "correction" to the mass as $\delta m := \frac{1}{2}\hat{K}^{(2,-1)}$.

Order $(2,0,\ln)$

We next move to the order $\ln(s)/s^2$ terms in the wave equation, and the order $\ln(s)/s$ terms in the gauge condition, which read

$$\ln s\,\partial^c\partial_c h^{(2,0,\ln)}_{B\alpha\beta} = 0, \tag{22.61}$$

$$\ln s\left(\partial^b h^{(2,0,\ln)}_{B\alpha b} - \tfrac{1}{2}\eta^{\mu\nu}x^a_\alpha\partial_a h^{(2,0,\ln)}_{B\mu\nu}\right) = 0. \tag{22.62}$$

From this we determine

$$h^{(2,0,\ln)}_{B\alpha\beta} = \hat{A}^{(2,0,\ln)}t_\alpha t_\beta + 2\hat{C}^{(2,0,\ln)}_a t_{(\beta}x^a_{\alpha)} + (\delta_{ab}\hat{K}^{(2,0,\ln)} + \hat{H}^{(2,0,\ln)}_{ab})x^a_\alpha x^b_\beta. \tag{22.63}$$

Finally, we arrive at the order $1/s^2$ terms in the wave equation. At this order, the body's tidal moments become coupled to those of the external background. The equation reads

$$\partial^c\partial_c h^{(2,0)}_{B\alpha\beta} + \frac{1}{s^2}\Big(h^{(2,0,\ln)}_{B\alpha\beta} + \tilde{E}_{\alpha\beta}\Big) = \frac{2}{s^2}\delta^2 R^{(0,-2)}_{\alpha\beta}\Big[h^{(1)}_B\Big], \tag{22.64}$$

where $\tilde{E}_{\alpha\beta}$ comprises the contributions from $h^{(2,-2)}_{B\alpha\beta}$ and $h^{(2,-1)}_{B\alpha\beta}$, given in Eqs. (A.10), (A.15), and (A.21). The contribution from the second-order Ricci tensor is given in Eqs. (A.7)–(A.9).

Foregoing the details, after some algebra we can read off the solution

$$h^{(2,0)}_{Btt} = \hat{A}^{(2,0)} + \hat{A}^{(2,0)}_i\omega^i + \hat{A}^{(2,0)}_{ij}\hat{\omega}^{ij} + \hat{A}^{(2,0)}_{ijk}\hat{\omega}^{ijk} \tag{22.65}$$

$$h^{(2,0)}_{Bta} = \hat{B}^{(2,0)}\omega_a + \hat{B}^{(2,0)}_{ij}\hat{\omega}_a{}^{ij} + \hat{C}^{(2,0)}_a + \hat{C}^{(2,0)}_{ai}\hat{\omega}_a{}^i + \epsilon_a{}^{bc}\left(\hat{D}^{(2,0)}_c\omega_b + \hat{D}^{(2,0)}_{ci}\hat{\omega}_b{}^i + \hat{D}^{(2,0)}_{cij}\hat{\omega}_b{}^{ij}\right) \tag{22.66}$$

$$\begin{aligned} h^{(2,0)}_{Bab} &= \delta_{ab}\left(\hat{K}^{(2,0)} + \hat{K}^{(2,0)}_i\omega^i + \hat{K}^{(2,0)}_{ijk}\hat{\omega}^{ijk}\right) + \hat{E}^{(2,0)}_i\hat{\omega}_{ab}{}^i + \hat{E}^{(2,0)}_{ij}\hat{\omega}_{ab}{}^{ij} + \hat{F}^{(2,0)}_{\langle a}\hat{\omega}_{b\rangle} + \hat{F}^{(2,0)}_{i\langle a}\hat{\omega}_{b\rangle}{}^i \\ &\quad + \hat{F}^{(2,0)}_{ij\langle a}\hat{\omega}_{b\rangle}{}^{ij} + \epsilon^{cd}{}_{(a}\hat{\omega}_{b)c}{}^i\hat{G}^{(2,0)}_{di} + \hat{H}^{(2,0)}_{ab} + \hat{H}^{(2,0)}_{abi}\omega^i + \epsilon^{cd}{}_{(a}\hat{I}^{(2,0)}_{b)d}\omega_c, \end{aligned} \tag{22.67}$$

where each one of the STF tensors is listed in Table 1.

In solving Eq. (22.64), we also find that the logarithmic term in the expansion becomes uniquely determined:

$$h^{(2,0,\ln)}_{B\alpha\beta} = -\tfrac{16}{15}m^2\mathcal{E}_{ab}x^a_\alpha x^b_\beta. \tag{22.68}$$

This term arises because the sources in the wave equation (22.64) contain a term $\propto \mathcal{E}_{ab}$, which cannot be equated to any term in $\partial^c\partial_c h^{(2,0)}_{Bab}$. Thus, the wave equation cannot be satisfied without including a logarithmic term.

Gauge condition

We now move to the final equation in the buffer region: the order $1/s$ gauge condition. This condition will determine the acceleration $a^\alpha_{(1)}$. At this order, $h^{(1)}_{\alpha\beta}$ first contributes to Eq. (22.19):

$$L^{(1,-1)}_\alpha[h^{(1)}] = \frac{4m}{s}a^{(1)}_a x^a_\alpha. \tag{22.69}$$

The contribution from $h^{(2)}_{B\alpha\beta}$ is most easily calculated by making use of Eqs. (B.24) and (B.25). After some algebra, we find that the t-component of the gauge condition reduces to

$$0 = -\frac{4}{s}\partial_t\delta m + \frac{4m}{3s}\partial_t\hat{A}^{(1,0)} + \frac{10m}{3s}\partial_t\hat{K}^{(1,0)}, \tag{22.70}$$

and the a-component reduces to

$$0 = \frac{4}{s}\partial_t^2 M_a + \frac{4m}{s}a^{(1)}_a + \frac{4}{s}\mathcal{E}_{ai}M^i + \frac{4}{s}\mathcal{B}_{ai}S^i - \frac{2m}{s}\hat{A}^{(1,1)}_a + \frac{4m}{s}\partial_t\hat{C}^{(1,0)}_a. \tag{22.71}$$

After removing common factors, these equations become Eqs. (22.39) and (22.41). We remind the reader that these equations are valid only when evaluated at $a^a(t) = a^a_{(0)}(t) = 0$, except in the term $4ma^{(1)}_a/s$ that arose from $L^{(1)}_\alpha[h^{(1)}]$.

Second-order solution

We have now completed our calculation in the buffer region. In summary, the second-order perturbation in the buffer region is given by $h^{(2)}_{\alpha\beta} = s^{-2}h^{(2,-2)}_{B\alpha\beta} + s^{-1}h^{(2,-1)}_{B\alpha\beta} + h^{(2,0)}_{B\alpha\beta} + \ln(s)h^{(2,0,\ln)}_{B\alpha\beta} + O(\varepsilon, s)$, where $h^{(2,-2)}_{B\alpha\beta}$ is given in Eq. (22.50), $h^{(2,-1)}_{B\alpha\beta}$ in Eq. (22.60), $h^{(2,0)}_{B\alpha\beta}$ in Eq. (22.65), and $h^{(2,0,\ln)}_{B\alpha\beta}$ in Eq. (22.68). In addition, we have found evolution equations for an effective correction to the body's mass, given by Eq. (22.39), and mass and spin dipoles, given by Eqs. (22.59) and (22.71).

Table 1: Symmetric trace-free tensors appearing in the order $\varepsilon^2 s^0$ part of the metric perturbation in the buffer region around the body. Each tensor is a function of the proper time t on the world line γ, and each is STF with respect to the Euclidean metric δ_{ij}.

$\hat{A}^{(2,0)}$		is arbitrary
$\hat{A}_i^{(2,0)}$	$=$	$-\partial_t^2 M_i - \frac{4}{5}S^j\mathcal{B}_{ji} + \frac{1}{3}M^j\mathcal{E}_{ji} - \frac{7}{5}m\hat{A}_i^{(1,1)} - \frac{3}{5}m\hat{K}_i^{(1,1)} + \frac{4}{5}m\partial_t\hat{C}_i^{(1,0)}$
$\hat{A}_{ij}^{(2,0)}$	$=$	$-\frac{7}{3}m^2\mathcal{E}_{ij}$
$\hat{A}_{ijk}^{(2,0)}$	$=$	$-2S_{\langle i}\mathcal{B}_{jk\rangle} + \frac{5}{3}M_{\langle i}\mathcal{E}_{jk\rangle} - \frac{1}{2}m\hat{H}_{ijk}^{(1,1)}$
$\hat{B}^{(2,0)}$	$=$	$m\partial_t\hat{K}^{(1,0)}$
$\hat{B}_{ij}^{(2,0)}$	$=$	$\frac{1}{9}\left(2M^l\mathcal{B}^k_{(i} - 5S^l\mathcal{E}^k_{(i}\right)\epsilon_{j)kl} - \frac{1}{2}m\hat{C}_{ij}^{(1,1)}$
$\hat{C}_i^{(2,0)}$		is arbitrary
$\hat{C}_{ij}^{(2,0)}$	$=$	$2\left(S^l\mathcal{E}^k_{(i} - \frac{14}{15}M^l\mathcal{B}^k_{(i}\right)\epsilon_{j)lk} - m\left(\frac{6}{5}\hat{C}_{ij}^{(1,1)} - \partial_t\hat{H}_{ij}^{(1,0)}\right)$
$\hat{D}_i^{(2,0)}$	$=$	$\frac{1}{5}\left(6M^j\mathcal{B}_{ij} - 7S^j\mathcal{E}_{ij}\right) + 2m\hat{D}_i^{(1,1)}$
$\hat{D}_{ij}^{(2,0)}$	$=$	$\frac{10}{3}m^2\mathcal{B}_{ij}$
$\hat{D}_{ijk}^{(2,0)}$	$=$	$\frac{1}{3}S_{\langle i}\mathcal{E}_{jk\rangle} + \frac{2}{3}M_{\langle i}\mathcal{B}_{jk\rangle}$
$\hat{K}^{(2,0)}$	$=$	$2\delta m$
$\hat{K}_i^{(2,0)}$	$=$	$-\partial_t^2 M_i - \frac{4}{5}S^j\mathcal{B}_{ij} - \frac{5}{9}M^j\mathcal{E}_{ij} + \frac{13}{15}m\hat{A}_i^{(1,1)} + \frac{9}{5}m\hat{K}_i^{(1,1)} - \frac{16}{15}m\partial_t\hat{C}_i^{(1,0)}$
$\hat{K}_{ijk}^{(2,0)}$	$=$	$-\frac{5}{9}M_{\langle i}\mathcal{E}_{jk\rangle} + \frac{2}{9}S_{\langle i}\mathcal{B}_{jk\rangle} - \frac{1}{6}m\hat{H}_{ijk}^{(1,1)}$
$\hat{E}_i^{(2,0)}$	$=$	$\frac{2}{15}M^i\mathcal{E}_{ij} + \frac{1}{5}S^j\mathcal{B}_{ij} + \frac{1}{10}m\partial_t\hat{C}_i^{(1,0)} - \frac{9}{20}m\hat{K}_i^{(1,1)} - \frac{11}{20}m\hat{A}_i^{(1,1)}$
$\hat{E}_{ij}^{(2,0)}$	$=$	$\frac{7}{5}m^2\mathcal{E}_{ij}$
$\hat{F}_i^{(2,0)}$	$=$	$\frac{184}{75}M^j\mathcal{E}_{ij} + \frac{72}{25}S^j\mathcal{B}_{ij} + \frac{46}{25}m\partial_t\hat{C}_i^{(1,0)} - \frac{28}{25}m\hat{A}_i^{(1,1)} + \frac{18}{25}m\hat{K}_i^{(1,1)}$
$\hat{F}_{ij}^{(2,0)}$	$=$	$4m^2\mathcal{E}_{ij}$
$\hat{F}_{ijk}^{(2,0)}$	$=$	$\frac{4}{3}M_{\langle i}\mathcal{E}_{jk\rangle} - \frac{4}{3}S_{\langle i}\mathcal{B}_{jk\rangle} + m\hat{H}_{ijk}^{(1,1)}$
$\hat{G}_{ij}^{(2,0)}$	$=$	$-\frac{4}{9}\epsilon_{lk(i}\mathcal{E}^k_{j)}M^l - \frac{2}{9}\epsilon_{lk(i}\mathcal{B}^k_{j)}S^l + \frac{1}{2}m\hat{I}_{ij}^{(1,1)}$
$\hat{H}_{ij}^{(2,0)}$		is arbitrary
$\hat{H}_{ijk}^{(2,0)}$	$=$	$\frac{58}{15}M_{\langle i}\mathcal{E}_{jk\rangle} - \frac{28}{15}S_{\langle i}\mathcal{B}_{jk\rangle} + \frac{2}{5}m\hat{H}_{ijk}^{(1,1)}$
$\hat{I}_{ij}^{(2,0)}$	$=$	$-\frac{104}{45}\epsilon_{lk(i}\mathcal{E}^k_{j)}M^l - \frac{112}{45}\epsilon_{lk(i}\mathcal{B}^k_{j)}S^l + \frac{8}{5}m\hat{I}_{ij}^{(1,1)}$

22.5 The equation of motion

Master equation of motion

Equation (22.71) is the principal result of our calculation. After simplification, it reads

$$\partial_t^2 M_a + \mathcal{E}_{ab}M^b = -a_a^{(1)} + \tfrac{1}{2}\hat{A}_a^{(1,1)} - \partial_t\hat{C}_a^{(1,0)} - \frac{1}{m}S_i\mathcal{B}_a^i. \tag{22.72}$$

We recall that M_a is the body's mass dipole moment, $\mathcal{E}_{ab}$ and $\mathcal{B}_{ab}$ are components of the Riemann tensor of the background spacetime evaluated on the world line, $a_a^{(1)}$ is the first-order acceleration of the world line, S_a is the body's spin angular momentum, and $\hat{A}_a^{(1,1)}$, $\hat{C}_a^{(1,0)}$ are vector fields on the world line that have yet to be determined. The equation is formulated in Fermi normal coordinates.

Equation (22.72) is a type of master equation of motion, describing the position of the body relative to a world line of unspecified (though small) acceleration, in terms of the metric perturbation on the world line, the tidal fields of the spacetime it lies in, and the spin of the body. It contains two types of accelerations: $\partial_t^2 M_i$ and $a^i_{(1)}$. The first type is the second time derivative of the body's mass dipole moment (or the first derivative of its ADM linear momentum), as measured in a frame centered on the world line γ. The second type is the covariant acceleration of the world line through the external spacetime. In other words, $\partial_t^2 M_i$ measures the acceleration of the body's centre of mass relative to the centre of the coordinate system, while a_i measures the acceleration of the coordinate system itself. As discussed in Sec. 21, our aim is to identify the world line as that of the body, and we do so via the condition that the mass dipole vanishes for all

times, meaning that the body is centered on the world line for all times. If we start with initial conditions $M_i(0) = 0 = \partial_t M_i(0)$, then the mass dipole remains zero for all times if and only if the world line satisfies the equation

$$a_a^{(1)} = \tfrac{1}{2}\hat{A}_a^{(1,1)} - \partial_t \hat{C}_a^{(1,0)} - \frac{1}{m} S_i \mathcal{B}_a^i. \tag{22.73}$$

This equation of motion contains two types of terms: a Papapetrou spin force, given by $-S_i\mathcal{B}_a^i$, which arises due to the coupling of the body's spin to the local magnetic-type tidal field of the external spacetime; and a self-force, arising from homogenous terms in the wave equation. Note that the right-hand side of this equation is to be evaluated at $a^\mu = a^{(0)\mu} = 0$, and that it would contain an antidamping term $-\frac{11}{3}m\dot{a}^\mu$ [7,149,150] if we had not assumed that the acceleration possesses an expansion of the form given in Eq. (21.9).

In our self-consistent approach, we began with the aim of identifying γ by the condition that the body must be centered about it for all time. However, we could have begun with a regular expansion, in which the world line is taken to be the remnant $\gamma^{(0)}$ of the body in the outer limit of $\varepsilon \to 0$ with only x^μ fixed. In that case the acceleration of the world line would necessarily be ε-independent, so $a^i_{(0)}$ would be the full acceleration of $\gamma^{(0)}$. Hence, when we found $a^i_{(0)} = 0$, we would have identified the world line as a geodesic, and there would be no corrections $a^i_{(n)}$ for $n > 0$. We would then have arrived at the equation of motion

$$\partial_t^2 M_a + \mathcal{E}_{ab}M^b = \tfrac{1}{2}\hat{A}_a^{(1,1)} - \partial_t \hat{C}_a^{(1,0)} - \frac{1}{m} S_i \mathcal{B}_a^i. \tag{22.74}$$

This equation of motion was first derived by Gralla and Wald [16] (although they phrased their expansion in terms of an explicit expansion of the world line, with a deviation vector on $\gamma^{(0)}$, rather than the mass dipole, measuring the correction to the motion). It describes the drift of the body away from the reference geodesic $\gamma^{(0)}$. This drift is driven partially by the local curvature of the background, as seen in the geodesic-deviation term $\mathcal{E}_{ab}M^b$, and by the coupling between the body's spin and the local curvature. It is also driven by the self-force, as seen in the terms containing $\hat{A}_a^{(1,1)}$ and $\hat{C}_a^{(1,0)}$, but unlike in the self-consistent equation, the fields that produce the self-force are generated by a geodesic past history (plus free propagation from initial data) rather than by the corrected motion.

Although perfectly valid, such an equation is of limited use. If the external background is curved, then M_i has meaning only if the body is "close" to the world line. Thus, $\partial_t^2 M_i$ is a meaningful acceleration only for a short time, since M_i will generically grow large as the body drifts away from the reference world line. On that short timescale of validity, the deviation vector defined by M^i accurately points from $\gamma^{(0)}$ to a "corrected" world line γ; that world line, the approximate equation of motion of which is given in Eq. (22.73), accurately tracks the motion of the body. After a short time, when the mass dipole grows large and the regular expansion scheme begins to break down, the deviation vector will no longer correctly point to the corrected world line. Errors will also accumulate in the field itself, because it is being sourced by the geodesic, rather than corrected, motion.

The self-consistent equation of motion appears to be more robust, and offers a much wider range of validity. Furthermore, even beyond the above step, where we had the option to choose to set either M_i or $a^i_{(0)}$ to zero, the self-consistent expansion continues to contain within it the regular expansion. Starting from the solution in the self-consistent expansion, one can recover the regular expansion, and its equation of motion (22.74), simply by assuming an expansion for the world line and following the usual steps of deriving the geodesic deviation equation.

Detweiler-Whiting decomposition

Regardless of which equation of motion we opt to use, we have now completed the derivation of the gravitational self-force, in the sense that, given the metric perturbation in the neighbourhood of the body, the self-force is uniquely determined by irreducible pieces of that perturbation. Explicitly, the terms that appear in the self-force are given by

$$\hat{A}_a^{(1,1)} = \frac{3}{4\pi}\int \omega_a h_{tt}^{(1,1)} d\Omega, \tag{22.75}$$

$$\hat{C}_a^{(1,0)} = h_{ta}^{(1,0)}. \tag{22.76}$$

This is all that is needed to incorporate the motion of the body into a dynamical system that can be numerically evolved; at each timestep, one simply needs to calculate the field near the world line and decompose it into irreducible pieces in order to determine the acceleration of the body. The remaining difficulty is to actually determine the field at each timestep. In the next section, we will use the formal integral representation of the solution to determine the metric perturbation at the location of the body in terms of a tail integral.

However, before doing so, we emphasize some important features of the self-force and the field near the body. First, note that the first-order external field $h^{(1)}_{\alpha\beta}$ splits into two distinct pieces. There is the singular piece $h^{\rm S}_{\alpha\beta}$, given by

$$h^{\rm S}_{tt} = \frac{2m}{s}\Big\{1 + \tfrac{3}{2}sa_i\omega^i + 2s^2a_ia^i + s^2\left(\tfrac{3}{8}a_{\langle i}a_{j\rangle} + \tfrac{5}{6}\mathcal{E}_{ij}\right)\hat{\omega}^{ij}\Big\} + O(s^2) \tag{22.77}$$

$$h^{\rm S}_{ta} = -2ms\dot{a}_a + \tfrac{2}{3}ms\epsilon_{aij}\mathcal{B}^j_k\hat{\omega}^{ik} + O(s^2) \tag{22.78}$$

$$\begin{aligned} h^{\rm S}_{ab} = \frac{2m}{s}\Big\{&\delta_{ab}\big[1 - \tfrac{1}{2}sa_i\omega^i + \tfrac{2}{3}s^2a_ia^i + s^2\left(\tfrac{3}{8}a_{\langle i}a_{j\rangle} - \tfrac{5}{18}\mathcal{E}_{ij}\right)\hat{\omega}^{ij}\big] + 2s^2a_{\langle a}a_{b\rangle} \\ &- \tfrac{19}{9}s^2\mathcal{E}_{ab} + \tfrac{2}{3}s^2\mathcal{E}^i_{\langle a}\hat{\omega}_{b\rangle i}\Big\} + O(s^2). \end{aligned} \tag{22.79}$$

This field is a solution to the homogenous wave equation for $s > 0$, but it is divergent at $s = 0$. It is the generalization of the $1/s$ Newtonian field of the body, as perturbed by the tidal fields of the external spacetime $g_{\alpha\beta}$. Comparing with results to be derived below in Sec. 23.2, we find that it is precisely the Detweiler-Whiting singular field for a point mass.

Next, there is the Detweiler-Whiting regular field $h^{\rm R}_{\alpha\beta} = h^{(1)}_{\alpha\beta} - h^{\rm S}_{\alpha\beta}$, given by

$$h^{\rm R}_{tt} = \hat{A}^{(1,0)} + s\hat{A}^{(1,1)}_i\omega^i + O(s^2), \tag{22.80}$$

$$h^{\rm R}_{ta} = \hat{C}^{(1,0)}_a + s\Big(\hat{B}^{(1,1)}\omega_a + \hat{C}^{(1,1)}_{ai}\omega^i + \epsilon_{ai}{}^j\hat{D}^{(1,1)}_j\omega^i\Big) + O(s^2), \tag{22.81}$$

$$h^{\rm R}_{ab} = \delta_{ab}\hat{K}^{(1,0)} + \hat{H}^{(1,0)}_{ab} + s\Big(\delta_{ab}\hat{K}^{(1,1)}_i\omega^i + \hat{H}^{(1,1)}_{abi}\omega^i + \epsilon_i{}^j{}_{(a}\hat{I}^{(1,1)}_{b)j}\omega^i + \hat{F}^{(1,1)}_{\langle a}\omega_{b\rangle}\Big) + O(s^2). \tag{22.82}$$

This field is a solution to the homogeneous wave equation even at $s = 0$. It is a free radiation field in the neighbourhood of the body. And it contains all the free functions in the buffer-region expansion.

Now, the acceleration of the body is given by

$$a^{(1)}_a = \tfrac{1}{2}\partial_a h^{\rm R}_{tt} - \partial_t h^{\rm R}_{ta} - \frac{1}{m}S_i\mathcal{B}^i_a, \tag{22.83}$$

which we can rewrite as

$$a^\alpha_{(1)} = -\tfrac{1}{2}\left(g^{\alpha\delta} + u^\alpha u^\delta\right)\left(2h^{\rm R}_{\delta\beta;\gamma} - h^{\rm R}_{\beta\gamma;\delta}\right)\big|_{a=0}u^\beta u^\gamma + \frac{1}{2m}R^\alpha{}_{\beta\gamma\delta}u^\beta S^{\gamma\delta}, \tag{22.84}$$

where $S^{\gamma\delta} := e^\gamma_c e^\delta_d \epsilon^{cdj}S_j$. In other words, up to order ε^2 errors, a body with order ε or smaller spin (i.e., one for which $S^{\gamma\delta} = 0$), moves on a geodesic of a spacetime $g_{\alpha\beta} + \varepsilon h^{\rm R}_{\alpha\beta}$, where $h^{\rm R}_{\alpha\beta}$ is a free radiation field in the neighbourhood of the body; a local observer would measure the "background spacetime," in which the body is in free fall, to have the metric $g_{\alpha\beta} + \varepsilon h^R_{\alpha\beta}$ instead of $g_{\alpha\beta}$. If we performed a transformation into Fermi coordinates in $g_{\alpha\beta} + \varepsilon h^R_{\alpha\beta}$, the metric would contain no acceleration term, and it would take the simple form of a smooth background plus a singular perturbation. Hence, the Detweiler-Whiting axiom is a consequence, rather than an assumption, of our derivation, and we have recovered precisely the picture it provides in the point particle case. In the electromagnetic and scalar cases, Harte has shown that this result is quite general: even for a finite extended body, the field it produces can be split into a homogeneous field [157,158,181] that exerts a direct force on the body, and a nonhomogeneous field that exerts only an indirect force by altering the body's multipole moments. His results should be generalizable to the gravitational case as well.

22.6 The effect of a gauge transformation on the force

We now turn to the question of how the world line transforms under a gauge transformation. We begin with the equation of motion (22.41), presented again here:

$$\partial_t^2 M_a + \mathcal{E}_{ai} M^i = -a_a^{(1)} - \frac{1}{m}\mathcal{B}_{ai} S^i + \left[\tfrac{1}{2}\hat{A}_a^{(1,1)} - \partial_t \hat{C}_a^{(1,0)}\right]_{a^\mu=0}. \tag{22.85}$$

Setting $M_i = 0$, we derive the first-order acceleration of γ, given in Eq. (22.73). If, for simplicity, we neglect the Papapetrou spin term, then that acceleration is given by

$$\begin{aligned} a_a^{(1)} &= \lim_{s\to 0}\left(\frac{3}{4\pi}\int \frac{\omega_a}{2s} h_{tt}^{(1)} d\Omega - \partial_t h_{ta}^{(1)}\right) \\ &= \lim_{s\to 0}\frac{3}{4\pi}\int \left(\tfrac{1}{2}\partial_i h_{tt}^{(1)} - \partial_t h_{ti}^{(1)}\right)\omega^i \omega_a d\Omega, \end{aligned} \tag{22.86}$$

where it is understood that explicit appearances of the acceleration are to be set to zero on the right-hand side. The first equality follows directly from Eq. (22.73) and the definitions of $\hat{A}_a^{(1,1)}$ and $\hat{C}_a^{(1,0)}$. The second equality follows from the STF decomposition of $h_{\alpha\beta}^{(1)}$ and the integral identities (B.26)–(B.28). We could also readily derive the form of the force given by the Quinn-Wald method of regularization: $\lim_{s\to 0}\frac{1}{4\pi}\int\left(\frac{1}{2}\partial_a h_{tt}^{(1)} - \partial_t h_{ta}^{(1)}\right)d\Omega$. However, in order to derive a gauge-invariant equation of motion, we shall use the form in Eq. (22.86).

Suppose that we had not chosen a world line for which the mass dipole vanishes, but instead had chosen some "nearby" world line. Then Eq. (22.85) provides the relationship between the acceleration of that world line, the mass dipole relative to it, and the first-order metric perturbations (we again neglect spin for simplicity). The mass dipole is given by $M_i = \frac{3}{8\pi}\lim_{s\to 0}\int s^2 h_{tt}^{(2)}\omega_i d\Omega$, which has the covariant form

$$M_{\alpha'} = \frac{3}{8\pi}\lim_{s\to 0}\int g^{\alpha}_{\alpha'}\omega_\alpha s^2 h_{\mu\nu}^{(2)} u^\mu u^\nu d\Omega, \tag{22.87}$$

where a primed index corresponds to a point on the world line. Note that the parallel propagator does not interfere with the angle-averaging, because in Fermi coordinates, $g^{\alpha}_{\beta'} = \delta^{\alpha}_{\beta} + O(\varepsilon, s^2)$. One can also rewrite the first-order-metric-perturbation terms in Eq. (22.85) using the form given in Eq. (22.86). We then have Eq. (22.85) in the covariant form

$$\begin{aligned} &\frac{3}{8\pi}\lim_{s\to 0}\int g^{\alpha}_{\alpha'}\left(g_{\alpha\beta}\frac{D^2}{d\tau^2} + \mathcal{E}_{\alpha\beta}\right)\omega^\beta s^2 h_{\mu\nu}^{(2)} u^\mu u^\nu d\Omega\big|_{a=a^{(0)}} \\ &= -\frac{3m}{8\pi}\lim_{s\to 0}\int g^{\alpha}_{\alpha'}\left(2h^{(1)}_{\beta\mu;\nu} - h^{(1)}_{\mu\nu;\beta}\right)u^\mu u^\nu \omega^\beta_\alpha d\Omega\big|_{a=a^{(0)}} - m a^{(1)}_{\alpha'}. \end{aligned} \tag{22.88}$$

Now consider a gauge transformation generated by $\varepsilon\xi^{(1)\alpha}[\gamma] + \frac{1}{2}\varepsilon^2\xi^{(2)\alpha}[\gamma] + \cdots$, where $\xi^{(1)\alpha}$ is bounded as $s\to 0$, and $\xi^{(2)\alpha}$ diverges as $1/s$. More specifically, we assume the expansions $\xi^{(1)\alpha} = \xi^{(1,0)\alpha}(t,\theta^A) + O(s)$ and $\xi^{(1)\alpha} = \frac{1}{s}\xi^{(2,-1)\alpha}(t,\theta^A) + O(1)$. (The dependence on γ appears in the form of dependence on proper time t. Other dependences could appear, but it would not affect the result.) This transformation preserves the presumed form of the outer expansion, both in powers of ε and in powers of s. The metric perturbations transform as

$$h^{(1)}_{\mu\nu} \to h^{(1)}_{\mu\nu} + 2\xi^{(1)}_{(\mu;\nu)}, \tag{22.89}$$

$$h^{(2)}_{\mu\nu} \to h^{(2)}_{\mu\nu} + \xi^{(2)}_{(\mu;\nu)} + h^{(1)}_{\mu\nu;\rho}\xi^{\rho}_{(1)} + 2h^{(1)}_{\rho(\mu}\xi^{\rho}_{(1);\nu)} + \xi^{\rho}_{(1)}\xi^{(1)}_{(\mu;\nu)\rho} + \xi^{\rho}_{(1);\mu}\xi^{(1)}_{\rho;\nu} + \xi^{\rho}_{(1);(\mu}\xi^{(1)}_{\nu);\rho}. \tag{22.90}$$

Using the results for $h^{(1)}_{\alpha\beta}$, the effect of this transformation on $h^{(2)}_{tt}$ is given by

$$h^{(2)}_{tt} \to h^{(2)}_{tt} - \frac{2m}{s^2}\omega^i\xi^{(1)}_i + O(s^{-1}). \tag{22.91}$$

The order $1/s^2$ term arises from $h^{(1)}_{\mu\nu;\rho}\xi^{\rho}_{(1)}$ in the gauge transformation. On the right-hand side of Eq. (22.88), the metric-perturbation terms transform as

$$(2h^{(1)}_{\beta\mu;\nu} - h^{(1)}_{\mu\nu;\beta})u^{\mu}u^{\nu}\omega^{\beta} \to (2h^{(1)}_{\beta\mu;\nu} - h^{(1)}_{\mu\nu;\beta})u^{\mu}u^{\nu}\omega^{\beta} + 2\omega_{\beta}\left(g^{\gamma}_{\beta}\frac{D^2}{d\tau^2} + \mathcal{E}^{\gamma}_{\beta}\right)\xi^{(1)}_{\gamma}. \tag{22.92}$$

The only remaining term in the equation is $ma^{\alpha}_{(1)}$. If we extend the acceleration off the world line in any smooth manner, then it defines a vector field that transforms as $a_{\alpha} \to a_{\alpha} + \varepsilon\pounds_{\xi_{(1)}}a_{\alpha} + \cdots$. Since $a^{\alpha}_{(0)} = 0$, this means that $a^{\alpha}_{(1)} \to a^{\alpha}_{(1)}$ — it is invariant under a gauge transformation. It is important to note that this statement applies to the acceleration on the original world line; it does not imply that the acceleration of the body itself is gauge-invariant.

From these results, we find that the left- and right-hand sides of Eq. (22.88) transform in the same way:

$$\text{LHS/RHS} \to \text{LHS/RHS} - \frac{3}{4\pi}\lim_{s\to 0}\int g^{\alpha}_{\alpha'}\omega^{\beta}_{\alpha}\left(g^{\gamma}_{\beta}\frac{D^2}{d\tau^2} + \mathcal{E}^{\gamma}_{\beta}\right)\xi^{(1)}_{\gamma}d\Omega. \tag{22.93}$$

Therefore, Eq. (22.88) provides a gauge-invariant relationship between the acceleration of a chosen fixed world line, the mass dipole of the body relative to that world line, and the first-order metric perturbations. So suppose that we begin in the Lorenz gauge, and we choose the fixed world line γ such that the mass dipole vanishes relative to it. Then in some other gauge, the mass dipole will no longer vanish relative to γ, and we must adopt a different, nearby fixed world line γ'. If the mass dipole is to vanish relative to γ', then the acceleration of that new world line must be given by $a^{\alpha} = \varepsilon a^{\alpha}_{(1)} + o(\varepsilon)$, where

$$a^{(1)}_{\alpha'} = -\frac{3m}{8\pi}\lim_{s\to 0}\int g^{\alpha}_{\alpha'}(2h^{(1)}_{\beta\mu;\nu} - h^{(1)}_{\mu\nu;\beta})u^{\mu}u^{\nu}\omega^{\beta}_{\alpha}d\Omega.\big|_{a=a^{(0)}}. \tag{22.94}$$

Hence, this is a covariant and gauge-invariant form of the first-order acceleration. By that we mean the *equation* is valid in any gauge, not that the value of the acceleration is the same in every gauge. Under a gauge transformation, a new fixed world line is adopted, and the value of the acceleration on it is related to that on the old world line according to Eq. (22.93). In the particular case that $\xi^{(1)}_{\mu}$ has no angle-dependence on the world line, this relationship reduces to

$$a^{(1)\alpha}_{\text{new}} = a^{(1)\alpha}_{\text{old}} - \left(g^{\alpha}_{\beta} + u^{\alpha}u_{\beta}\right)\left(\frac{D^2\xi^{\beta}_{(1)}}{d\tau^2} + R^{\beta}{}_{\mu\nu\rho}u^{\mu}\xi^{\nu}_{(1)}u^{\rho}\right), \tag{22.95}$$

as first derived by Barack and Ori [19]. (Here we've replaced the tidal field with its expression in terms of the Riemann tensor to more transparently agree with equations in the literature.) An argument of this form was first presented by Gralla [182] for the case of a regular expansion, and was extended to the case of a self-consistent expansion in Ref. [18].

23 Global solution in the external spacetime

In the previous sections, we have determined the equation of motion of γ in terms of the metric perturbation; we now complete the first-order solution by determining the metric perturbation. In early derivations of the gravitational self-force (excluding those in Refs. [16, 183]), the first-order external perturbation was simply assumed to be that of a point particle. This was first justified by Gralla and Wald [16]. An earlier argument made by D'Eath [164, 165] (and later used by Rosenthal [184]) provided partial justification but was incomplete [17]. Here, we follow the derivation in Ref. [17], which makes use of the same essential elements as D'Eath's: the integral formulation of the perturbative Einstein equation and the asymptotically small radius of the tube Γ.

23.1 Integral representation

Suppose we take our buffer-region expansion of $h^{(1)}_{\alpha\beta}$ to be valid everywhere in the interior of Γ (in $\mathcal{M}_E$), rather than just in the buffer region. This is a meaningful supposition in a distributional sense, since the

$1/s$ singularity in $h^{(1)}_{\alpha\beta}$ is locally integrable even at γ. Note that the extension of the buffer-region expansion is not intended to provide an accurate or meaningful approximation in the interior; it is used only as a means of determining the field in the exterior. We can do this because the field values in Ω are entirely determined by the field values on Γ, so using the buffer-region expansion in the interior of Γ leaves the field values in Ω unaltered. Now, given the extension of the buffer-region expansion, it follows from Stokes' law that the integral over Γ in Eq. (21.8) can be replaced by a volume integral over the interior of the tube, plus two surface integrals over the "caps" $\mathcal{J}_{\rm cap}$ and $\Sigma_{\rm cap}$, which fill the "holes" in $\mathcal{J}$ and Σ, respectively, where they intersect Γ. Schematically, we can write Stokes' law as $\int_{\rm Int(\Gamma)} = \int_{\mathcal{J}_{\rm cap}} + \int_{\Sigma_{\rm cap}} - \int_{\Gamma}$, where Int($\Gamma$) is the interior of Γ. This is now valid as a distributional identity. (Note that the "interior" here means the region bounded by $\Gamma \cup \Sigma_{\rm cap} \cup \mathcal{J}_{\rm cap}$; Int($\Gamma$) does not refer to the set of interior points in the point-set defined by Γ.) The minus sign in front of the integral over Γ accounts for the fact that the directed surface element in Eq. (21.8) points *into* the tube. Because $\mathcal{J}_{\rm cap}$ does not lie in the past of any point in Ω, it does not contribute to the perturbation at $x \in \Omega$. Hence, we can rewrite Eq. (21.8) as

$$\begin{aligned} h^{(1)}_{\alpha\beta} &= -\frac{1}{4\pi}\int\limits_{\rm Int(\Gamma)} \nabla_{\mu'}\left(G^{+}_{\alpha\beta}{}^{\alpha'\beta'}\nabla^{\mu'}h^{(1)}_{\alpha'\beta'} - h^{(1)}_{\alpha'\beta'}\nabla^{\mu'}G^{+}_{\alpha\beta}{}^{\alpha'\beta'}\right)dV' + h^{(1)}_{\bar\Sigma\alpha\beta} \\ &= -\frac{1}{4\pi}\int\limits_{\rm Int(\Gamma)}\left(G_{\alpha\beta}{}^{\alpha'\beta'}E_{\alpha'\beta'}[h^{(1)}] - h^{(1)}_{\alpha'\beta'}E^{\alpha'\beta'}[G^{+}_{\alpha\beta}]\right)dV' + h^{(1)}_{\bar\Sigma\alpha\beta}, \end{aligned} \tag{23.1}$$

where $h^{(1)}_{\bar\Sigma\alpha\beta}$ is the contribution from the spatial surface $\bar\Sigma := \Sigma \cup \Sigma_{\rm cap}$, and $E^{\alpha'\beta'}[G^{+}_{\alpha\beta}]$ denotes the action of the wave-operator on $G^{+}_{\alpha\beta}{}^{\gamma'\delta'}$. Now note that $E^{\alpha'\beta'}[G^{+}_{\alpha\beta}] \propto \delta(x,x')$; since $x \notin$ Int(Γ), this term integrates to zero. Next note that $E_{\alpha'\beta'}[h^{(1)}_{\alpha\beta}]$ vanishes everywhere except at γ. This means that the field at x can be written as

$$h^{(1)}_{\alpha\beta} = \frac{-1}{4\pi}\lim_{\mathcal{R}\to 0}\int\limits_{\rm Int(\Gamma)} G^{+}_{\alpha\beta}{}^{\alpha'\beta'}E_{\alpha'\beta'}[h^{(1)}]dV' + h^{(1)}_{\bar\Sigma\alpha\beta}. \tag{23.2}$$

Making use of the fact that $E_{\alpha\beta}[h^{(1)}] = \partial^c\partial_c(1/s)h^{(1,-1)}_{\alpha\beta} + O(s^{-2})$, along with the identity $\partial^c\partial_c(1/s) = -4\pi\delta^3(x^a)$, where δ^3 is a coordinate delta function in Fermi coordinates, we arrive at the desired result

$$h^{(1)}_{\alpha\beta} = 2m\int_\gamma G^{+}_{\alpha\beta\bar\alpha\bar\beta}\left(2u^{\bar\alpha}u^{\bar\beta} + g^{\bar\alpha\bar\beta}\right)d\bar t + h^{(1)}_{\bar\Sigma\alpha\beta}. \tag{23.3}$$

Therefore, in the region Ω, the leading-order perturbation produced by the asymptotically small body is identical to the field produced by a point particle. At second order, the same method can be used to simplify Eq. (21.5) by replacing at least part of the integral over Γ with an integral over γ. We will not pursue this simplification here, however.

Gralla and Wald [16] provided an alternative derivation of the same result, using distributional methods to prove that the distributional source for the linearized Einstein equation must be that of a point particle in order for the solution to diverge as $1/s$. One can understand this by considering that the most divergent term in the linearized Einstein tensor is a Laplacian acting on the perturbation, and the Laplacian of $1/s$ is a flat-space delta function; the less divergent corrections are due to the curvature of the background, which distorts the flat-space distribution into a covariant curved-spacetime distribution.

23.2 Metric perturbation in Fermi coordinates

Metric perturbation

We have just seen that the solution to the wave equation with a point-mass source is given by

$$h^{(1)}_{\alpha\beta} = 2m\int_\gamma G_{\alpha\beta\alpha'\beta'}(2u^{\alpha'}u^{\beta'} + g^{\alpha'\beta'})dt' + h^{(1)}_{\Sigma\alpha\beta}. \tag{23.4}$$

One can also obtain this result from Eq. (19.29) by taking the trace-reversal and making use of the Green's function identity (16.22). In this section, we seek an expansion of the perturbation in Fermi coordinates. Following the same steps as in Sec. 19.2, we arrive at

$$h^{(1)}_{\alpha\beta} = \frac{2m}{r} U_{\alpha\beta\alpha'\beta'}(2u^{\alpha'}u^{\beta'} + g^{\alpha'\beta'}) + h^{\rm tail}_{\alpha\beta}(u). \tag{23.5}$$

Here primed indices now refer to the retarded point $z^\alpha(u)$ on the world line, r is the retarded radial coordinate at x, and the tail integral is given by

$$\begin{aligned} h^{\rm tail}_{\alpha\beta}(u) &= 2m\int_{t^<}^{u} V_{\alpha\beta\alpha'\beta'}(2u^{\alpha'}u^{\beta'} + g^{\alpha'\beta'})dt' + 2m\int_0^{t^<} G_{\alpha\beta\alpha'\beta'}(2u^{\alpha'}u^{\beta'} + g^{\alpha'\beta'})dt' + h^{(1)}_{\Sigma\alpha\beta} \\ &= 2m\int_0^{u^-} G_{\alpha\beta\alpha'\beta'}(2u^{\alpha'}u^{\beta'} + g^{\alpha'\beta'})dt' + h^{(1)}_{\Sigma\alpha\beta}, \end{aligned} \tag{23.6}$$

where $t^<$ is the first time at which the world line enters $\mathcal{N}(x)$, and $t = 0$ denotes the time when it crosses the initial-data surface Σ.

We expand the direct term in $h^{(1)}_{\alpha\beta}$ in powers of s using the following: the near-coincidence expansion $U_{\alpha\beta}{}^{\alpha'\beta'} = g^{\alpha'}_{\alpha} g^{\beta'}_{\beta}(1 + O(s^3))$; the relationship between r and s, given by Eq. (11.5); and the coordinate expansion of the parallel-propagators, obtained from the formula $g^{\alpha'}_{\alpha} = u^{\alpha'}e^0_\alpha + e^{\alpha'}_a e^a_\alpha$, where the retarded tetrad (u^α, e^α_a) can be expanded in terms of s using Eqs. (11.9), (11.10), (9.12), and (9.13). We expand the tail integral similarly: noting that $u = t-s+O(s^2)$, we expand $h^{\rm tail}_{\alpha\beta}(u)$ about t as $h^{\rm tail}_{\alpha\beta}(t) - s\partial_t h^{\rm tail}_{\alpha\beta}(t) + O(s^2)$; each term is then expanded using the near-coincidence expansions $V_{\alpha\beta}{}^{\alpha''\beta''} = g^{\gamma''}_{(\alpha} g^{\delta''}_{\beta)} R^{\alpha''}{}_{\gamma''}{}^{\beta''}{}_{\delta''} + O(s)$ and $h^{\rm tail}_{\alpha\beta}(t) = g^{\bar\alpha}_{\alpha} g^{\bar\beta}_{\beta}(h^{\rm tail}_{\bar\alpha\bar\beta} + s h^{\rm tail}_{\bar\alpha\bar\beta i}\omega^i) + O(s^2)$, where barred indices correspond to the point $\bar x = z(t)$, and $h^{\rm tail}_{\bar\alpha\bar\beta\bar\gamma}$ is given by

$$h^{\rm tail}_{\bar\alpha\bar\beta\bar\gamma} = 2m\int_0^{t^-} \nabla_{\bar\gamma} G_{\bar\alpha\bar\beta\alpha'\beta'}(2u^{\alpha'}u^{\beta'} + g^{\alpha'\beta'})dt' + h^{(1)}_{\Sigma\bar\alpha\bar\beta\bar\gamma}. \tag{23.7}$$

This yields the expansion

$$h^{\rm tail}_{\alpha\beta}(u) = g^{\bar\alpha}_{\alpha} g^{\bar\beta}_{\beta}(h^{\rm tail}_{\bar\alpha\bar\beta} + s h^{\rm tail}_{\bar\alpha\bar\beta i}\omega^i - 4ms\mathcal{E}_{\bar\alpha\bar\beta}) + O(s^2). \tag{23.8}$$

As with the direct part, the final coordinate expansion is found by expressing $g^{\bar\alpha}_{\alpha}$ in terms of the Fermi tetrad.

Combining the expansions of the direct and tail parts of the perturbation, we arrive at the expansion in Fermi coordinates:

$$\begin{aligned} h^{(1)}_{tt} &= \frac{2m}{s}\left(1 + \tfrac{3}{2}sa_i\omega^i + \tfrac{3}{8}s^2 a_i a_j \omega^{ij} - \tfrac{15}{8}s^2\dot a_{\bar\alpha}u^{\bar\alpha} + \tfrac{1}{3}s^2\dot a_i\omega^i + \tfrac{5}{6}s^2\mathcal{E}_{ij}\omega^{ij}\right) + (1 + 2sa_i\omega^i)h^{\rm tail}_{00} \\ &\quad + s h^{\rm tail}_{00i}\omega^i + O(s^2), \end{aligned} \tag{23.9}$$

$$h^{(1)}_{ta} = 4ma_a - \tfrac{2}{3}msR_{0iaj}\omega^{ij} + 2ms\mathcal{E}_{ai}\omega^i - 2ms\dot a_a + (1 + sa_i\omega^i)h^{\rm tail}_{0a} + s h^{\rm tail}_{0ai}\omega^i + O(s^2), \tag{23.10}$$

$$\begin{aligned} h^{(1)}_{ab} &= \frac{2m}{s}\left(1 - \tfrac{1}{2}sa_i\omega^i + \tfrac{3}{8}s^2 a_i a_j \omega^{ij} + \tfrac{1}{8}s^2\dot a_{\bar\alpha}u^{\bar\alpha} + \tfrac{1}{3}s^2\dot a_i\omega^i - \tfrac{1}{6}s^2\mathcal{E}_{ij}\omega^{ij}\right)\delta_{ab} + 4msa_a a_b \\ &\quad - \tfrac{2}{3}msR_{aibj}\omega^{ij} - 4ms\mathcal{E}_{ab} + h^{\rm tail}_{ab} + s h^{\rm tail}_{abi}\omega^i + O(s^2). \end{aligned} \tag{23.11}$$

As the final step, each of these terms is decomposed into irreducible STF pieces using the formulas (B.1), (B.3), and (B.7), yielding

$$h^{(1)}_{tt} = \frac{2m}{s} + \hat A^{(1,0)} + 3ma_i\omega^i + s\left[4ma_i a^i + \hat A^{(1,1)}_i\omega^i + m\left(\tfrac{3}{4}a_{\langle i}a_{j\rangle} + \tfrac{5}{3}\mathcal{E}_{ij}\right)\hat\omega^{ij}\right] + O(s^2), \tag{23.12}$$

$$h^{(1)}_{ta} = \hat C^{(1,0)}_a + s\left(\hat B^{(1,1)}\omega_a - 2m\dot a_a + \hat C^{(1,1)}_{ai}\omega^i + \epsilon_{ai}{}^{j}\hat D^{(1,1)}_j\omega^i + \tfrac{2}{3}m\epsilon_{aij}\mathcal{B}^j_k\hat\omega^{ik}\right) + O(s^2), \tag{23.13}$$

$$\begin{aligned} h^{(1)}_{ab} &= \frac{2m}{s}\delta_{ab} + (\hat K^{(1,0)} - ma_i\omega^i)\delta_{ab} + \hat H^{(1,0)}_{ab} + s\Big\{\delta_{ab}\left[\tfrac{4}{3}ma_i a^i + \hat K^{(1,1)}_i\omega^i + \tfrac{3}{4}ma_{\langle i}a_{j\rangle}\hat\omega^{ij} - \tfrac{5}{9}m\mathcal{E}_{ij}\hat\omega^{ij}\right] \\ &\quad + \tfrac{4}{3}m\mathcal{E}^i_{\langle a}\hat\omega_{b\rangle i} + 4ma_{\langle a}a_{b\rangle} - \tfrac{38}{9}m\mathcal{E}_{ab} + \hat H^{(1,1)}_{abi}\omega^i + \epsilon_i{}^j{}_{(a}\hat I^{(1,1)}_{b)j}\omega^i + \hat F^{(1,1)}_{\langle a}\omega_{b\rangle}\Big\} + O(s^2), \end{aligned} \tag{23.14}$$

Table 2: Symmetric trace-free tensors in the first-order metric perturbation in the buffer region, written in terms of the electric-type tidal field $\mathcal{E}_{ab}$, the acceleration a_i, and the tail of the perturbation.

$\hat{A}^{(1,0)} = h^{\rm tail}_{00}$	$\hat{C}^{(1,1)}_{ab} = h^{\rm tail}_{0\langle ab\rangle} + 2m\mathcal{E}_{ab} + h^{\rm tail}_{0\langle a}a_{b\rangle}$
$\hat{C}^{(1,0)}_a = h^{\rm tail}_{0a} + ma_a$	$\hat{D}^{(1,1)}_a = \frac{1}{2}\epsilon_a{}^{bc}(h^{\rm tail}_{0bc} + h^{\rm tail}_{0b}a_c)$
$\hat{K}^{(1,0)} = \frac{1}{3}\delta^{ab}h^{\rm tail}_{ab}$	$\hat{K}^{(1,1)}_a = \frac{1}{3}\delta^{bc}h^{\rm tail}_{bca} + \frac{2}{3}m\dot{a}_a$
$\hat{H}^{(1,0)}_{ab} = h^{\rm tail}_{\langle ab\rangle}$	$\hat{H}^{(1,1)}_{abc} = h^{\rm tail}_{\langle abc\rangle}$
$\hat{A}^{(1,1)}_a = h^{\rm tail}_{00a} + 2h^{\rm tail}_{00}a_a + \frac{2}{3}m\dot{a}_a$	$\hat{F}^{(1,1)}_a = \frac{3}{5}\delta^{ij}h^{\rm tail}_{\langle ia\rangle j}$
$\hat{B}^{(1,1)} = \frac{1}{3}h^{\rm tail}_{0ij}\delta^{ij} + \frac{1}{3}h^{\rm tail}_{0i}a^i$	$\hat{I}^{(1,1)}_{ab} = \frac{2}{3}\mathop{\rm STF}_{ab}\left(\epsilon_b{}^{ij}h^{\rm tail}_{\langle ai\rangle j}\right)$

where the uppercase hatted tensors are specified in Table 2. Because the STF decomposition is unique, these tensors must be identical to the free functions in Eq. (22.38); hence, those free functions, comprising a regular, homogenous solution in the buffer region, have been uniquely determined by boundary conditions and waves emitted by the particle in the past.

Singular and regular pieces

The Detweiler-Whiting singular field is given by

$$h^{\rm S}_{\alpha\beta} = 2m\int G^S_{\alpha\beta\alpha'\beta'}(2u^{\alpha'}u^{\beta'} + g^{\alpha'\beta'})dt'. \tag{23.15}$$

Using the Hadamard decomposition $G^S_{\alpha\beta\alpha'\beta'} = \frac{1}{2}U_{\alpha\beta\alpha'\beta'}\delta(\sigma) - \frac{1}{2}V_{\alpha\beta\alpha'\beta'}\theta(\sigma)$, we can write this as

$$\begin{aligned} h^{\rm S}_{\alpha\beta} &= \frac{m}{r}U_{\alpha\beta\alpha'\beta'}(2u^{\alpha'}u^{\beta'} + g^{\alpha'\beta'}) + \frac{m}{r_{\rm adv}}U_{\alpha\beta\alpha''\beta''}(2u^{\alpha''}u^{\beta''} + g^{\alpha''\beta''}) \\ &\quad - 2m\int_u^v V_{\alpha\beta\bar{\alpha}\bar{\beta}}(u^{\bar{\alpha}}u^{\bar{\beta}} + \tfrac{1}{2}g^{\bar{\alpha}\bar{\beta}})d\bar{t}, \end{aligned} \tag{23.16}$$

where primed indices now refer to the retarded point $x' = z(u)$; double-primed indices refer to the advanced point $x'' = z(v)$; barred indices refer to points in the segment of the world line between $z(u)$ and $z(v)$. The first term in Eq. (23.16) can be read off from the calculation of the retarded field. The other terms are expanded using the identities $v = u + 2s + O(s^2)$ and $r_{\rm adv} = r(1 + \frac{2}{3}s^2\dot{a}_i\omega^i)$. The final result is

$$h^{\rm S}_{tt} = \frac{2m}{s} + 3ma_i\omega^i + ms\big[4a_ia^i + \tfrac{3}{4}a_{\langle i}a_{j\rangle}\hat{\omega}^{ij} + \tfrac{5}{3}\mathcal{E}_{ij}\hat{\omega}^{ij}\big] + O(s^2), \tag{23.17}$$

$$h^{\rm S}_{ta} = s\big(-2m\dot{a}_a + \tfrac{2}{3}m\epsilon_{aij}\mathcal{B}^j_k\hat{\omega}^{ik}\big) + O(s^2), \tag{23.18}$$

$$\begin{aligned} h^{\rm S}_{ab} &= \frac{2m}{s}\delta_{ab} - ma_i\omega^i\delta_{ab} + s\Big\{\delta_{ab}\big[\tfrac{4}{3}ma_ia^i + \big(\tfrac{3}{4}ma_{\langle i}a_{j\rangle} - \tfrac{5}{9}m\mathcal{E}_{ij}\big)\hat{\omega}^{ij}\big] + \tfrac{4}{3}m\mathcal{E}^i_{\langle a}\hat{\omega}_{b\rangle i} \\ &\quad + 4ma_{\langle a}a_{b\rangle} - \tfrac{38}{9}m\mathcal{E}_{ab}\Big\} + O(s^2). \end{aligned} \tag{23.19}$$

The regular field could be calculated from the regular Green's function. But it is more straightforwardly calculated using $h^{\rm R}_{\alpha\beta} = h^{(1)}_{\alpha\beta} - h^{\rm S}_{\alpha\beta}$. The result is

$$h^{\rm R}_{tt} = \hat{A}^{(1,0)} + s\hat{A}^{(1,1)}_i\omega^i + O(s^2), \tag{23.20}$$

$$h^{\rm R}_{ta} = \hat{C}^{(1,0)}_a + s\Big(\hat{B}^{(1,1)}\omega_a + \hat{C}^{(1,1)}_{ai}\omega^i + \epsilon_{ai}{}^j\hat{D}^{(1,1)}_j\omega^i\Big) + O(s^2), \tag{23.21}$$

$$h^{\rm R}_{ab} = \delta_{ab}\hat{K}^{(1,0)} + \hat{H}^{(1,0)}_{ab} + s\Big(\delta_{ab}\hat{K}^{(1,1)}_i\omega^i + \hat{H}^{(1,1)}_{abi}\omega^i + \epsilon_i{}^j{}_{(a}\hat{I}^{(1,1)}_{b)j}\omega^i + \hat{F}^{(1,1)}_{\langle a}\omega_{b\rangle}\Big) + O(s^2). \tag{23.22}$$

23.3 Equation of motion

With the metric perturbation fully determined, we can now express the self-force in terms of tail integrals. Reading off the components of $h^{\rm R}_{\alpha\beta}$ from Table 2 and inserting the results into Eq. (22.84), we arrive at

$$a^\mu_{(1)} = -\frac{1}{2}\left(g^{\mu\nu} + u^\mu u^\nu\right)\left(2h^{\rm tail}_{\nu\lambda\rho} - h^{\rm tail}_{\lambda\rho\nu}\right)\Big|_{a=0} u^\lambda u^\rho + \frac{1}{2m}R^\mu{}_{\nu\lambda\rho}u^\nu S^{\lambda\rho}. \tag{23.23}$$

We have now firmly established the results of the point-particle analysis.

24 Concluding remarks

We have presented a number of derivations of the equations that determine the motion of a point scalar charge q, a point electric charge e, and a point mass m in a specified background spacetime. In this concluding section we summarize these derivations and their foundations. We conclude by describing the next step in the gravitational case: obtaining an approximation scheme sufficiently accurate to extract the parameters of an extreme-mass-ratio inspiral from an observed gravitational waveform.

Our derivations are of two types. The first is based on the notion of an exact point particle. In this approach, we assume that the self-force on the particle arises from a particular piece of its field, either that which survives angle-averaging or the Detweiler-Whiting regular field. The second type is based on the notion of an asymptotically small body, and abandons the fiction of a point particle. In this approach, we don't assume anything about the body's equation of motion, but rather derive it directly from the field equations. Although we have presented such a derivation only in the gravitational case, analogous ones could be performed in the scalar and electromagnetic cases, using conservation of energy-momentum instead of the field equations alone. Such a calculation was performed by Gralla *et al.* in the restricted case of an electric charge in a flat background [2].

Perhaps the essential result of our derivation based on an asymptotically small body is that it confirms all of the results derived using point particles: at linear order in the body's mass, the field it creates is identical to that of a point particle, and its equation of motion is precisely that derived from physically motivated axioms for a point particle. In other words, at linear order, not only can we get away with the fiction of a point particle, but our assumptions about the physics governing its motion are also essentially correct.

24.1 The motion of a point particle

Spatial averaging

Our first means of deriving equations of motion for point particles is based on spatial averaging. In this approach, we assume the following axiom:

the force on the particle arises from the piece of the field that survives angle averaging.

For convenience in our review, we consider the case of a point electric charge and adopt the Detweiler-Whiting decomposition of the Faraday tensor into singular and regular pieces, $F_{\alpha\beta} = F^{\rm R}_{\alpha\beta} + F^{\rm S}_{\alpha\beta}$. We average $F_{\alpha\beta}$ over a sphere of constant proper distance from the particle. We then evaluate the averaged field at the particle's position. Because the regular field is nonsingular on the world line, this yields

$$e\langle F_{\mu\nu}\rangle u^\nu = e\langle F^{\rm S}_{\mu\nu}\rangle u^\nu + eF^{\rm R}_{\mu\nu}u^\nu,$$

where

$$e\langle F^{\rm S}_{\mu\nu}\rangle u^\nu = -(\delta m)a_\mu, \qquad \delta m = \lim_{s\to 0}\left(\frac{2}{3}\frac{e^2}{s}\right),$$

and

$$eF^{\rm R}_{\mu\nu}u^\nu = e^2\big(g_{\mu\nu} + u_\mu u_\nu\big)\left(\frac{2}{3}\dot{a}^\nu + \frac{1}{3}R^\nu{}_\lambda u^\lambda\right) + 2e^2u^\nu\int_{-\infty}^{\tau^-}\nabla_{[\mu}G^+{}_{\nu]\lambda'}\big(z(\tau),z(\tau')\big)u^{\lambda'}\,d\tau'.$$

We now postulate that the equations of motion are $ma_\mu = e\langle F_{\mu\nu}\rangle u^\nu$, where m is the particle's bare mass. With the preceding results we arrive at $m_{\rm obs}a_\mu = eF^{\rm R}_{\mu\nu}u^\nu$, where $m_{\rm obs} \equiv m + \delta m$ is the particle's observed (renormalized) inertial mass.

In this approach, the fiction of a point particle manifests itself in the need for mass renormalization. Such a requirement can be removed, even within the point-particle picture, by adopting the "comparison axiom" proposed by Quinn and Wald [7]. If we consider extended (but small) bodies, no such renormalization is required, and the equations of motion follow directly from the conservation of energy-momentum. However, the essential assumption about the nature of the force is valid: only the piece of the field that survives angle-averaging exerts a force on the body.

The Detweiler-Whiting Axiom

Our second means of deriving equations of motion for point particles is based on the Detweiler-Whiting axiom, which asserts that

> the singular field exerts no force on the particle; the entire self-force arises from the action of the regular field.

This axiom, which is motivated by the symmetric nature of the singular field, and also its causal structure, gives rise to the same equations of motion as the averaging method. In this picture, the particle simply interacts with a free field (whose origin can be traced to the particle's past), and the procedure of mass renormalization is sidestepped. In the scalar and electromagnetic cases, the picture of a particle interacting with a free radiation field removes any tension between the nongeodesic motion of the charge and the principle of equivalence. In the gravitational case the Detweiler-Whiting axiom produces a generalized equivalence principle (c.f. Ref. [156]): up to order ε^2 errors, a point mass m moves on a geodesic of the spacetime with metric $g_{\alpha\beta} + h^{\mathrm{R}}_{\alpha\beta}$, which is nonsingular and a solution to the vacuum field equations. This is a conceptually powerful, and elegant, formulation of the MiSaTaQuWa equations of motion. And it remains valid for (non-spinning) small bodies.

Resolving historical ambiguities

Although they yield the correct physical description, the above axioms are by themselves insufficient, and historically, two problems have arisen in utilizing them: One, they led to ill-behaved equations of motion, requiring a process of order reduction; and two, in the gravitational case they led to equations of motion that are inconsistent with the field equations, requiring the procedure of gauge-relaxation. Both of these problems arose because the expansions were insufficiently systematic, in the sense that they did not yield exactly solvable perturbation equations. In the approach taken in our review, we have shown that these problems do not arise within the context of a systematic expansion. Although we have done so only in the case of an extended body, where we sought a higher degree of rigor, one could do the same in the case of point particles by expanding in the limit of small charge or mass (see, e.g., the treatment of a point mass in Ref. [17]).

Consider the Abraham-Lorentz-Dirac equation $ma^\mu = f^\mu_{\mathrm{ext}} + \frac{2}{3}e^2\dot{a}^\mu$. To be physically meaningful and mathematically well-justified, it must be thought of as an approximate equation of motion for a localized matter distribution with small charge $e \ll 1$. But it contains terms of differing orders e^0 and e^2, and the acceleration itself is obviously a function of e. Hence, the equation has not been fully expanded. One might think that it is somehow an exact equation, despite its ill behaviour. Or one might replace it with the order-reduced equation $ma^\mu = f^\mu_{\mathrm{ext}} + \frac{2e^2}{3m}\dot{f}^\mu_{\mathrm{ext}}$ to eliminate that ill behaviour. But one can instead assume that $a^\mu(e)$, like other functions of e, possesses an expansion in powers of e, leading to the two well-behaved equations $ma^{(0)\mu} = f^\mu_{\mathrm{ext}}$ and $ma^{(1)\mu} = \frac{2}{3}e^2\dot{a}^{(0)\mu}$. However, the fact that such an equation can even arise indicates that one has not begun with a systematic expansion of the governing field equations (in this case, the conservation equation and the Maxwell equations). If one began with a systematic expansion, with equations exactly solvable at each order, no such ambiguity would arise.

The same can be said of the second problem. It is well known that in general relativity, the motion of gravitating bodies is determined by the Einstein field equations; the equations of motion cannot be separately imposed. And specifically, if we deal with the linearized Einstein equation $\delta G_{\alpha\beta}[h^{(1)}] = 8\pi T_{\alpha\beta}[\gamma]$, where $T_{\alpha\beta}$ is the energy-momentum tensor of a point particle in the background spacetime, then the linearized Bianchi identity requires the point particle to move on a geodesic of the background spacetime. This seems to contradict the MiSaTaQuWa equation and therefore the assumptions we made in deriving it. In order

to remove this inconsistency, the earliest derivations [6, 7] invoked an a posteriori gauge-relaxation: rather than solving a linearized Einstein equation exactly, they solved the wave equation $E_{\alpha\beta}[h^{(1)}] = -16\pi T_{\alpha\beta}$ in combination with the relaxed gauge condition $L_\alpha[\varepsilon h^{(1)}] = O(\varepsilon^2)$. The allowed errors in the gauge condition carry over into the linearized Bianchi identity, such that it no longer restricts the motion to be geodesic. In this approach, one is *almost* solving the linearized problem one set out to solve. But again, such an a posteriori corrective measure is required only if one begins without a systematic expansion. The first-order metric perturbation is a functional of a world line; if we allow that world line to depend on ε, then the metric has evidently not been fully expanded in powers of ε. To resolve the problem, one needs to carefully deal with this fact.

In the approach we have adopted here, following Ref. [17], we have resolved these problems via a self-consistent expansion in which the world line is held fixed while expanding the metric. Rather than beginning with the linear field equation, we began by reformulating the exact equation in the form

$$E_{\alpha\beta}[h] = T^{\text{eff}}_{\alpha\beta}, \tag{24.1}$$

$$L_\alpha[h] = 0. \tag{24.2}$$

These equations we systematically expanded by (i) treating the metric perturbation as a functional of a fixed world line, keeping the dependence on the world line fixed while expanding, and (ii) expanding the acceleration of the world line. Since we never sought a solution to the equation $\delta G_{\alpha\beta}[h^{(1)}] = 8\pi T_{\alpha\beta}[\gamma]$, no tension arose between the equation of motion and the field equation. In addition, we dealt only with exactly solvable perturbative equations: rather then imposing the ad hoc a posteriori gauge condition $L_\mu[\varepsilon h^{(1)}] = O(\varepsilon^2)$, our approach systematically led to the conditions $L^{(0)}_\mu\left[h^{(1)}\right] = 0$ and $L^{(1)}_\mu\left[h^{(1)}\right] = -L^{(0)}_\mu\left[h^{(2)}\right]$, which can be solved exactly. Like in the issue of order reduction, the essential step in arriving at exactly solvable equations is assuming an expansion of the particle's acceleration.

Historically, these issues were first resolved by Gralla and Wald [16] using a different method. Rather than allowing an ε-dependence in the first-order perturbation, they fully expanded every function in the problem in a power series, including the world line itself. In this approach, the world line of the body is found to be a geodesic, but higher-order effects arise in deviation vectors measuring the drift of the particle away from that geodesic. Such a method has the drawback of being limited to short timescales, since the deviation vectors will eventually grow large as the body moves away from the initial reference geodesic.

24.2 The motion of a small body

Although the above results for point particles require assumptions about the form of the force, their results have since been derived from first principles, and the physical pictures they are based on have proven to be justified: An asymptotically small body behaves as a point particle moving on a geodesic of the smooth part of the spacetime around it, or equivalently, it moves on a world line accelerated by the asymmetric part of its own field.

In addition to a derivation from first principles, we also seek a useful approximation scheme. Any such scheme must deal with the presence of multiple distinct scales. Most obviously, these are the mass and size of the body itself and the lengthscales of the external universe, but other scales also arise. For example, in order to accurately represent the physics of an extreme-mass-ratio system, we must consider changes over four scales: First, there is the large body's mass, which is the characteristic lengthscale of the external universe. For convenience, since all lengths are measured relative to this one, we rescale them such that this global lengthscale is ~ 1. Second, there is the small body's mass $\sim \varepsilon$, which is the scale over which the gravitational field changes near the body; since the body is compact, this is also the scale of its linear size. Third, there is the radiation-reaction time $\sim 1/\varepsilon$; this is the time over which the small effects of the self-force accumulate to produce significant changes, specifically the time required for quantities such as the small body's energy and angular momentum to accumulate order 1 changes. Fourth, there is the large distance to the wave zone. We will not discuss this last scale here, but dealing with it analytically would likely require matching a wave solution at null infinity to an expansion formally expected to be valid in a region of size $\sim 1/\varepsilon$.

Self-consistent and matched asymptotic expansions

In this review, we have focused on a self-consistent approximation scheme first presented in Refs. [17,18,129]. It deals with the small size of the body using two expansions. Near the body, to capture changes on the short distances $\sim \varepsilon$, we adopt an inner expansion in which the body remains of constant size while all other distances approach infinity. Outside of this small neibourhood around the body, we adopt an outer expansion in which the body shrinks to zero mass and size about an ε-dependent world line that accurately reflects its long-term motion. The world line γ is defined in the background spacetime of the outer expansion. Its acceleration is found by solving the Einstein equation in a buffer region surrounding the body, where both expansions are valid; in this region, both expansions must agree, and we can use the multipole moments of the inner expansion to determine the outer one. In particular, we define γ to be the body's world line if and only if the body's mass dipole moment vanishes when calculated in coordinates centered on it. As in the point-particle calculation, the essential step in arriving at exactly solvable equations and a well-behaved equation of motion is an assumed expansion of the acceleration on the world line.

In order to construct a global solution in the outer expansion, we first recast the Einstein equation in a form that can be expanded and solved for an arbitrary world line. As in the point-particle case, we accomplish this by adopting the Lorenz gauge for the total metric perturbation. We can then write formal solutions to the perturbation equations at each order as integrals over a small tube surrounding the body (plus an initial data surface). By embedding the tube in the buffer region, we can use the data from the buffer-region expansion to determine the global solution. At first order, we find that the metric perturbation is precisely that of a point particle moving on the world line γ. While the choice of gauge is not essential in finding an expression for the force in terms of the field in the buffer-region expansion, it *is* essential in our method of determining the global field. Without making use of the relaxed Einstein equations, no clear method of globally solving the Einstein equation presents itself, and no simple split between the perturbation and the equation of motion arises.

Because this expansion self-consistently incorporates the corrections to the body's motion, it promises to be accurate on long timescales. Specifically, when combined with the first-order metric perturbation, the first-order equation of motion defines a solution to the Einstein equation that we expect to be accurate up to order ε^2 errors over times $t \lesssim 1/\varepsilon$. When combined with the second-order perturbation, it defines a solution we expect to be accurate up to order ε^3 errors on the shorter timescale ~ 1.

This approach closely mirrors the extremely successful methods of post-Newtonian theory [156]. In particular, both schemes recast the Einstein equation in a relaxed form before expanding it. Also, our use of an inner limit near the body is analogous to the "strong-field point particle limit" exploited by Futamase and his collaborators [156]. And our calculation of the equation of motion is somewhat similar to the methods used by Futamase and others [156, 174], in that it utilizes a multipole expansion of the body's metric in the buffer region.

Alternative methods

Various other approaches have been taken to deal with the multiple scales in the problem. In particular, even the earliest paper on the gravitational self-force [6] made use of inner and outer limits, which were also used in different forms in later derivations [15, 33, 183]. However, those early derivations are problematic. Specifically, they never adequately define the world line for which they derive equations of motion.

In the method of matched asymptotic expansions used in Refs. [6, 15, 33], the first-order perturbation in the outer expansion is assumed to be that of a point particle, with an attendant world line, and the inner expansion is assumed to be that of a perturbed black hole; by matching the two expansions in the buffer region, the acceleration of the world line is determined. Since the forms of the inner and outer expansions are already restricted, this approach's conclusions have somewhat limited strength, but it has more fundamental problems as well. The matching procedure begins by expanding the outer expansion in powers of distance s (or r) from the world line, and the inner expansion for large spacial distances $R \gg \varepsilon$. But the two expansions begin in different coordinate systems with an unknown relationship between them. In particular, there is no given relationship between the world line and the "position" of the black hole. The two expansions are matched by finding a coordinate transformation that makes them agree in the buffer region. However, because there is no predetermined relationship between the expansions, this transformation is not in fact unique, and it does not yield a unique equation of motion. One can increase the strength of the matching

condition in order to arrive at unique results, but that further weakens the strength of the conclusions. Refer to Ref. [18] for a discussion of these issues.

Rather than finding the equation of motion from the field equations, as in the above calculations, Fukumoto *et al.* [183] found the equation of motion by defining the body's linear momentum as an integral over the body's interior and then taking the derivative of that momentum. But they then required an assumed relationship between the momentum and the four-velocity of a representative world line in the body's interior. Hence, the problem is again an inadequate defintion of the body's motion. Because it involves integrals over the body's interior, and takes the world line to lie therein, this approach is also limited to material bodies; it does not apply to black holes.

The first reliable derivation of the first-order equation of motion for an asymptotically small body was performed by Gralla and Wald [16], who used a method of the same nature as the one presented here, deriving equations of motion by solving the field equations in the buffer region. As we have seen, however, their derivation is based on an expansion of the world line in powers of ε instead of a self-consistent treatment that keeps it fixed.

Using very different methods, Harte has also provided reliable derivations of equations of motion for extended bodies interacting with their own scalar and electromagnetic fields in fixed background spacetimes [157, 158, 181]. His approach is based on generalized definitions of momenta, the evolution of which is equivalent to energy-momentum conservation. The momenta are defined in terms of generalized Killing fields ξ^α, the essential property of which is that they satisfy $\pounds_\xi g_{\alpha\beta}|_\gamma = \nabla_\gamma \pounds_\xi g_{\alpha\beta}|_\gamma = 0$ — that is, they satisfy Killing's equation on the body's world line and approximately satisfy it "nearby." Here the world line can be defined in multiple ways using, for example, center-of-mass conditions. This approach is nonperturbative, with no expansion in the limit of small mass and size (though it does require an upper limit on the body's size and a lower limit on its compactness). It has the advantage of very naturally deriving a generalization of the Detweiler-Whiting axiom: The field of an extended body can be split into (i) a solution to the vacuum field equations which exerts a direct self-force, and (ii) a solution to the equations sourced by the body, which shifts the body's multipole moments. This approach has not yet been applied to gravity, but such an application should be relatively straightforward. However, as with the approach of Fukumoto *et al.* [183], this one does not apply to black holes.

Other methods have also been developed (or suggested) to accomplish the same goals as the self-consistent expansion. Most prominent among these is the two-timescale expansion suggested by Hinderer and Flanagan [121]. As discussed in Sec. 2, their method splits the orbital evolution into slow and fast dynamics by introducing slow and fast time variables. In the terminology of Sec. 22, this method constructs a general expansion by smoothly transitioning between regular expansions constructed at each value of the slow time variable, with the transition determined by the evolution with respect to the slow time. On the scale of the fast time, the world line is a geodesic; but when the slow time is allowed to vary, the world line transitions between geodesics to form the true, accelerated world line. This results in a global, uniform-in-time approximation. One should note that simply patching together a sequence of regular expansions, by shifting to a new geodesic every so often using the deviation vector, would not accomplish this: Such a procedure would accumulate a secular error in both the metric perturbation and the force, because the perturbation would be sourced by a world line secularly deviating from the position of the body, and the force would be calculated from this erroneous perturbation. The error would be proportional to the number of "shifts" multiplied by a nonlinear factor depending on the time between them. And this error would, formally at least, be of the same magnitude as the solution itself.

The fundamental difference between the self-consistent expansion and the two-timescale expansion is the following: In the two-timescale method, the Einstein equation, coupled to the equation of motion of the small body, is reduced to a dynamical system that can be evolved in time. The true world line of the body then emerges from the evolution of this system. In the method presented here, we have instead sought global, formal solutions to the Einstein equation, written in terms of global integrals; to accomplish this, we have treated the world line of the body as a fixed structure in the external spacetime. However, the two methods should agree. Note, though, that the two-timescale expansion is limited to orbits in Kerr, and it requires evolution equations for the slow evolution of the large black hole's mass and spin parameters, which have not yet been derived. Since the changes in mass and spin remain small on a radiation-reaction timescale, in the self-consistent expansion presented here they are automatically incorporated into the perturbations $h^{(n)}_{\alpha\beta}$. It is possible that this incorporation leads to errors on long timescales, in which case a different approach,

naturally allowing slow changes in the background, would be more advantageous.

24.3 Beyond first order

The primary experimental motivation for researching the self-force is to produce waveform templates for LISA. In order to extract the parameters of an extreme-mass-ratio binary from a waveform, we require a waveform that is accurate up to errors of order ε after a radiation-reaction time $\sim 1/\varepsilon$. If we use the first-order equation of motion, we will be neglecting an acceleration $\sim \varepsilon^2$, which will lead to secular errors of order unity after a time $\sim 1/\varepsilon$. Thus, the second-order self-force is required in order to obtain a sufficiently accurate waveform template. In order to achieve the correct waveform, we must also obtain the second-order part of the metric perturbation; this can be easily done, at least formally, using the global integral representations outside a world tube. However, a practical numerical calculation may prove difficult, since one would not wish to excise the small tube from one's numerical domain, and the second-order perturbation would diverge too rapidly on the world line to be treated straightforwardly.

A formal expression for the second-order force has already been derived by Rosenthal [184,185]. However, he expresses the second-order force in a very particular gauge in which the first-order self-force vanishes. This is sensible on short timescales, but not on long timescales, since it forces secular changes into the first-order perturbation, presumably leading to the first-order perturbation becoming large with time. Furthermore, it is not a convenient gauge, since it does not provide what we wish it to: a correction to the nonzero leading-order force in the Lorenz gauge.

Thus, we wish to obtain an alternative to Rosenthal's derivation. Based on the methods reviewed in this article, there is a clear route to deriving the second-order force. One would construct a buffer-region expansion accurate up to order ε^3. Since one would require the order $\varepsilon^2 s$ terms in this expansion, in order to determine the acceleration, one would need to increase the order of the expansion in r as well. Specifically, one would need terms up to orders $\varepsilon^0 s^3$, εs^2, $\varepsilon^2 s$, and $\varepsilon^3 s^0$. In such a calculation, one would expect the following terms to appear: the body's quadrupole moment Q_{ab}, corrections δM_i and δS_i to its mass and spin dipoles, and a second-order correction $\delta^2 m$ to its mass. Although some ambiguity may arise in defining the world line of the body at this order, a reasonable definition appears to be to guarantee that δM_i vanishes. However, at this order one may require some model of the body's internal dynamics, since the equation of motion will involve the body's quadrupole moment, for which the Einstein equation is not expected to yield an evolution equation. But if one seeks only the second-order self-force, one could simply neglect the quadrupole by assuming that the body is spherically symmetric in isolation. In any case, the force due to the body's quadrupole moment is already known from various other methods: see, e.g., the work of Dixon [159–161]; more recent methods can be found in Ref. [186] and references therein.

Because such a calculation could be egregiously lengthy, one may consider simpler methods, perhaps requiring stronger assumptions. For example, one could straightforwardly implement the method of matched asymptotic expansions with the matching conditions discussed in Ref. [18], in which one makes strong assumptions about the relationship between the inner and outer expansions. The effective field theory method used by Galley and Hu [163] offers another possible route.

Alternatively, one could calculate only part of the second-order force. Specifically, as described in Sec. 2.5, Hinderer and Flanagan [121, 130] have shown that one requires in fact only the averaged dissipative part of the second-order force. This piece of the force can be calculated within an adiabatic approximation, in which the rates of change of orbital parameters are calculated from the radiative Green's function, asymptotic wave amplitudes, and information about the orbit that sources them. Hence, we might be able to forgo a complete calculation of the second-order force, and use instead the complete first-order force in conjunction with an adiabatic approximation for the second-order force.

The self-force, however, is of interest beyond its relevance to LISA. The full second-order force would be useful for more general purposes, such as more accurate comparisons to post-Newtonian theory, and analysis of other systems, such as intermediate mass ratio binaries. Perhaps most importantly, it is of fundamental importance in our understanding of the motion of small bodies. For these reasons, proceeding to second order in a systematic expansion, and thereby obtaining second-order expressions for the force on a small body, remains an immediate goal.

Appendices

A Second-order expansions of the Ricci tensor

We present here various expansions used in solving the second-order Einstein equation in Sec. 22.4. We require an expansion of the second-order Ricci tensor $\delta^2 R_{\alpha\beta}$, defined by

$$\begin{aligned}\delta^2 R_{\alpha\beta}[h] = &-\tfrac{1}{2}\gamma^{\mu\nu}{}_{;\nu}\left(2h_{\mu(\alpha;\beta)} - h_{\alpha\beta;\mu}\right) + \tfrac{1}{4}h^{\mu\nu}{}_{;\alpha}h_{\mu\nu;\beta} + \tfrac{1}{2}h^{\mu}{}_{\beta}{}^{;\nu}\left(h_{\mu\alpha;\nu} - h_{\nu\alpha;\mu}\right)\\ &- \tfrac{1}{2}h^{\mu\nu}\left(2h_{\mu(\alpha;\beta)\nu} - h_{\alpha\beta;\mu\nu} - h_{\mu\nu;\alpha\beta}\right),\end{aligned} \tag{A.1}$$

where $\gamma^{\mu\nu}$ is the trace-reversed metric perturbation, and an expansion of a certain piece of $E_{\mu\nu}[h^{(2)}]$. Specifically, we require an expansion of $\delta^2 R^{(0)}_{\alpha\beta}[h^{(1)}]$ in powers of the Fermi radial coordinate s, where for a function f, $\delta^2 R^{(0)}_{\alpha\beta}[f]$ consists of $\delta^2 R_{\alpha\beta}[f]$ with the acceleration a^μ set to zero. We write

$$\delta^2 R^{(0)}_{\alpha\beta}[h^{(1)}] = \frac{1}{s^4}\delta^2 R^{(0,-4)}_{\alpha\beta}\left[h^{(1)}\right] + \frac{1}{s^3}\delta^2 R^{(0,-3)}_{\alpha\beta}\left[h^{(1)}\right] + \frac{1}{s^2}\delta^2 R^{(0,-2)}_{\alpha\beta}\left[h^{(1)}\right] + O(1/s), \tag{A.2}$$

where the second superscript index in parentheses denotes the power of s. Making use of the expansion of $h^{(1)}_{\alpha\beta}$, obtained by setting the acceleration to zero in the results for $h^{(1)}_{\alpha\beta}$ found in Sec. 22.3, one finds

$$\delta^2 R^{(2,-4)}_{\alpha\beta}\left[h^{(1)}\right] = 2m^2\left(7\hat{\omega}_{ab} + \tfrac{4}{3}\delta_{ab}\right)x^a_\alpha x^b_\beta - 2m^2 t_\alpha t_\beta, \tag{A.3}$$

and

$$\delta^2 R^{(2,-3)}_{tt}\left[h^{(1)}\right] = 3m\hat{H}^{(1,0)}_{ij}\hat{\omega}^{ij}, \tag{A.4}$$

$$\delta^2 R^{(2,-3)}_{ta}\left[h^{(1)}\right] = 3m\hat{C}^{(1,0)}_{i}\hat{\omega}^{i}_{a}, \tag{A.5}$$

$$\delta^2 R^{(2,-3)}_{ab}\left[h^{(1)}\right] = 3m\big(\hat{A}^{(1,0)} + \hat{K}^{(1,0)}\big)\hat{\omega}_{ab} - 6m\hat{H}^{(1,0)}_{i\langle a}\hat{\omega}^{i}_{b\rangle} + m\delta_{ab}\hat{H}^{(1,0)}_{ij}\hat{\omega}^{ij}, \tag{A.6}$$

and

$$\begin{aligned}\delta^2 R^{(2,-2)}_{tt}\left[h^{(1)}\right] = &-\tfrac{20}{3}m^2\mathcal{E}_{ij}\hat{\omega}^{ij} + 3m\hat{H}^{(1,1)}_{ijk}\hat{\omega}^{ijk} + \tfrac{7}{5}m\hat{A}^{(1,1)}_{i}\omega^i + \tfrac{3}{5}m\hat{K}^{(1,1)}_{i}\omega^i\\ &- \tfrac{4}{5}m\partial_t\hat{C}^{(1,0)}_{i}\omega^i,\end{aligned} \tag{A.7}$$

$$\begin{aligned}\delta^2 R^{(2,-2)}_{ta}\left[h^{(1)}\right] = &-m\partial_t\hat{K}^{(1,0)}\omega_a + 3m\hat{C}^{(1,1)}_{ij}\hat{\omega}_a{}^{ij} + m\Big(\tfrac{6}{5}\hat{C}^{(1,1)}_{ai} - \partial_t\hat{H}^{(1,0)}_{ai}\Big)\omega^i\\ &+ 2m\epsilon_a{}^{ij}\hat{D}^{(1,1)}_{i}\omega_j + \tfrac{4}{3}m^2\epsilon_{aik}\mathcal{B}^k_j\hat{\omega}^{ij},\end{aligned} \tag{A.8}$$

$$\begin{aligned}\delta^2 R^{(2,-2)}_{ab}\left[h^{(1)}\right] = &\,\delta_{ab}m\left(\tfrac{16}{15}\partial_t\hat{C}^{(1,0)}_{i} - \tfrac{13}{15}\hat{A}^{(1,1)}_{i} - \tfrac{9}{5}\hat{K}^{(1,1)}_{i}\right)\omega^i\\ &+ \delta_{ab}\left(-\tfrac{50}{9}m^2\mathcal{E}_{ij}\hat{\omega}^{ij} + m\hat{H}^{(1,1)}_{ijk}\hat{\omega}^{ijk}\right) - \tfrac{14}{3}m^2\mathcal{E}_{ij}\hat{\omega}_{ab}{}^{ij}\\ &+ m\left(\tfrac{33}{10}\hat{A}^{(1,1)}_{i} + \tfrac{27}{10}\hat{K}^{(1,1)}_{i} - \tfrac{3}{5}\partial_t\hat{C}^{(1,0)}_{i}\right)\hat{\omega}_{ab}{}^{i}\\ &+ m\left(\tfrac{28}{25}\hat{A}^{(1,1)}_{\langle a} - \tfrac{18}{25}\hat{K}^{(1,1)}_{\langle a} - \tfrac{46}{25}\partial_t\hat{C}^{(1,0)}_{\langle a}\right)\hat{\omega}_{b\rangle}\\ &- \tfrac{8}{3}m^2\mathcal{E}_{i\langle a}\hat{\omega}_{b\rangle}{}^{i} - 6m\hat{H}^{(1,1)}_{ij\langle a}\hat{\omega}_{b\rangle}{}^{ij} + 3m\epsilon_{ij(a}\hat{\omega}_{b)}{}^{jk}\hat{I}^{(1,1)}_{ik}\\ &+ \tfrac{2}{45}m^2\mathcal{E}_{ab} - \tfrac{2}{5}m\hat{H}^{(1,1)}_{abi}\omega^i + \tfrac{8}{5}m\epsilon^i{}_{j(a}\hat{I}^{(1,1)}_{b)i}\omega^j.\end{aligned} \tag{A.9}$$

Next, we require an analogous expansion of $E^{(0)}_{\alpha\beta}\left[\frac{1}{r^2}h^{(2,-2)} + \frac{1}{r}h^{(2,-1)}\right]$, where $E^{(0)}_{\alpha\beta}[f]$ is defined for any f by setting the acceleration to zero in $E_{\alpha\beta}[f]$. The coefficients of the $1/r^4$ and $1/r^3$ terms in this expansion

can be found in Sec. 22.4; the coefficient of $1/r^2$ will be given here. For compactness, we define this coefficient to be $\tilde{E}_{\alpha\beta}$. The tt-component of this quantity is given by

$$\tilde{E}_{tt} = 2\partial_t^2 M_i\omega^i + \tfrac{8}{5}S^j\mathcal{B}_{ij}\omega^i - \tfrac{2}{3}M^j\mathcal{E}_{ij}\omega^i + \tfrac{82}{3}m^2\mathcal{E}_{ij}\hat{\omega}^{ij} + 24S_{\langle i}\mathcal{B}_{jk\rangle}\hat{\omega}^{ijk} - 20M_{\langle i}\mathcal{E}_{jk\rangle}\hat{\omega}^{ijk}. \tag{A.10}$$

The ta-component is given by

$$\begin{aligned}\tilde{E}_{ta} &= \tfrac{44}{15}\epsilon_{aij}M^k\mathcal{B}^j_k\omega^i - \tfrac{2}{15}\left(11S^i\mathcal{E}^j_k + 18M^i\mathcal{B}^j\right)\epsilon_{ija}\omega^k + \tfrac{2}{15}\left(41S^j\mathcal{E}^k_a - 10M^j\mathcal{B}^k_a\right)\epsilon_{ijk}\omega^i \\ &\quad + 4\epsilon_{aij}\left(S^j\mathcal{E}_{kl} + 2M_k\mathcal{B}^j_l\right)\hat{\omega}^{ikl} + 4\epsilon_{ij\langle k}\mathcal{E}^j_{l\rangle}S^i\hat{\omega}_a{}^{kl} + \tfrac{68}{3}m^2\epsilon_{aij}\mathcal{B}^j_k\hat{\omega}^{ik}.\end{aligned} \tag{A.11}$$

This can be decomposed into irreducible STF pieces via the identities

$$\epsilon_{aij}S^i\mathcal{E}^j_k = S^i\mathcal{E}^j_{(k}\epsilon_{a)ij} + \tfrac{1}{2}\epsilon_{akj}S^i\mathcal{E}^j_i \tag{A.12}$$

$$\epsilon_{aj\langle i}\mathcal{E}_{kl\rangle}S^j = \underset{ikl}{\mathrm{STF}}\left[\epsilon^j{}_{al}S_{\langle i}\mathcal{E}_{jk\rangle} - \tfrac{2}{3}\delta_{al}S^p\mathcal{E}^j_{(i}\epsilon_{k)jp}\right] \tag{A.13}$$

$$\epsilon_{aj\langle i}M_l\mathcal{B}_{k\rangle}{}^j = \underset{ikl}{\mathrm{STF}}\left[\epsilon^j{}_{al}M_{\langle i}\mathcal{B}_{jk\rangle} + \tfrac{1}{3}\delta_{al}M^p\mathcal{B}^j_{(i}\epsilon_{k)jp}\right], \tag{A.14}$$

which follow from Eqs. (B.3) and (B.7), and which lead to

$$\begin{aligned}\tilde{E}_{ta} &= \tfrac{2}{5}\epsilon_{aij}\left(6M^k\mathcal{B}^j_k - 7S^k\mathcal{E}^j_k\right)\omega^i + \tfrac{4}{3}\left(2M^l\mathcal{B}^k_{(i} - 5S^l\mathcal{E}^k_{(i}\right)\epsilon_{j)kl}\hat{\omega}_a{}^{ij} \\ &\quad + \left(4S^j\mathcal{E}^k_{(a} - \tfrac{56}{15}M^j\mathcal{B}^k_{(a}\right)\epsilon_{i)jk}\omega^i + 4\epsilon_{ai}{}^l\left(S_{\langle j}\mathcal{E}_{kl\rangle} + 2M_{\langle j}\mathcal{B}_{kl\rangle}\right)\hat{\omega}^{ijk} + \tfrac{68}{3}m^2\epsilon_{aij}\mathcal{B}^j_k\hat{\omega}^{ik}.\end{aligned} \tag{A.15}$$

The ab-component is given by

$$\begin{aligned}\tilde{E}_{ab} &= \tfrac{56}{3}m^2\mathcal{E}_{ij}\hat{\omega}_{ab}{}^{ij} + \tfrac{52}{45}m^2\mathcal{E}_{ab} - \delta_{ab}\left[\left(2\partial_t^2M_i + \tfrac{8}{5}S^j\mathcal{B}_{ij} + \tfrac{10}{9}M^j\mathcal{E}_{ij}\right)\omega^i + \tfrac{100}{9}m^2\mathcal{E}_{ij}\hat{\omega}^{ij}\right] \\ &\quad - \delta_{ab}\left(\tfrac{20}{3}M_{\langle i}\mathcal{E}_{jk\rangle} - \tfrac{8}{3}S_{\langle i}\mathcal{B}_{jk\rangle}\right)\hat{\omega}^{ijk} + \tfrac{8}{15}M_{\langle a}\mathcal{E}_{b\rangle i}\omega^i + \tfrac{8}{15}M^i\mathcal{E}_{i\langle a}\omega_{b\rangle} + \tfrac{56}{3}m^2\mathcal{E}_{i\langle a}\hat{\omega}_{b\rangle}{}^i \\ &\quad + 16M_i\mathcal{E}_{j\langle a}\hat{\omega}_{b\rangle}{}^{ij} - \tfrac{32}{5}S_{\langle a}\mathcal{B}_{b\rangle i}\omega^i + \tfrac{4}{15}\left(10S_i\mathcal{B}_{ab} + 27M_i\mathcal{E}_{ab}\right)\omega^i \\ &\quad + \tfrac{16}{3}S^i\mathcal{B}_{i\langle a}\omega_{b\rangle} - 8\epsilon_{ij\langle a}\epsilon_{b\rangle kl}S^j\mathcal{B}^l_m\hat{\omega}^{ikm} + \tfrac{16}{15}\epsilon_{ij\langle a}\epsilon_{b\rangle kl}S^j\mathcal{B}^{il}\omega^k.\end{aligned} \tag{A.16}$$

Again, this can be decomposed, using the identities

$$S_{\langle a}\mathcal{B}_{b\rangle i} = S_{\langle a}\mathcal{B}_{bi\rangle} + \underset{ab}{\mathrm{STF}}\,\tfrac{1}{3}\epsilon_{ai}{}^j\epsilon_{kl(b}\mathcal{B}_{j)}{}^lS^k + \tfrac{1}{10}\delta_{i\langle a}\mathcal{B}_{b\rangle j}S^j, \tag{A.17}$$

$$S_i\mathcal{B}_{ab} = S_{\langle a}\mathcal{B}_{bi\rangle} - \underset{ab}{\mathrm{STF}}\,\tfrac{2}{3}\epsilon_{ai}{}^j\epsilon_{kl(b}\mathcal{B}_{j)}{}^lS^k + \tfrac{3}{5}\delta_{i\langle a}\mathcal{B}_{b\rangle j}S^j, \tag{A.18}$$

$$\epsilon_{ij\langle a}\epsilon_{b\rangle kl}S^j\mathcal{B}^{il} = \underset{ab}{\mathrm{STF}}\,\epsilon_{akj}S^l\mathcal{B}^i_{(j}\epsilon_{b)il} - \tfrac{1}{2}\delta_{k\langle a}\mathcal{B}_{b\rangle i}S^i, \tag{A.19}$$

$$\underset{ikm}{\mathrm{STF}}\,\epsilon_{ij\langle a}\epsilon_{b\rangle kl}S^j\mathcal{B}^l_m = \underset{ikm}{\mathrm{STF}}\,\underset{ab}{\mathrm{STF}}\left(2\delta_{ai}S_{\langle b}\mathcal{B}_{km\rangle} + \tfrac{1}{3}\delta_{ai}\epsilon^l{}_{bk}S^j\mathcal{B}^p_{(l}\epsilon_{m)jp} - \tfrac{3}{10}\delta_{ai}\delta_{bk}\mathcal{B}_{mj}S^j\right), \tag{A.20}$$

which lead to

$$\begin{aligned}\tilde{E}_{ab} &= -2\delta_{ab}\left[\left(\partial_t^2M_i + \tfrac{4}{5}S^j\mathcal{B}_{ij} + \tfrac{5}{9}M^j\mathcal{E}_{ij}\right)\omega^i + \left(\tfrac{10}{3}M_{\langle i}\mathcal{E}_{jk\rangle} - \tfrac{4}{3}S_{\langle i}\mathcal{B}_{jk\rangle}\right)\hat{\omega}^{ijk}\right] \\ &\quad - \tfrac{100}{9}\delta_{ab}m^2\mathcal{E}_{ij}\hat{\omega}^{ij} + \tfrac{1}{5}\left(8M^j\mathcal{E}_{ij} + 12S^j\mathcal{B}_{ij}\right)\hat{\omega}_{ab}{}^i + \tfrac{56}{3}m^2\mathcal{E}_{ij}\hat{\omega}_{ab}{}^{ij} \\ &\quad + \tfrac{4}{75}\left(92M^j\mathcal{E}_{j\langle a} + 108S^j\mathcal{B}_{j\langle a}\right)\omega_{b\rangle} + \tfrac{56}{3}m^2\mathcal{E}_{i\langle a}\hat{\omega}_{b\rangle}{}^i \\ &\quad + 16\,\underset{aij}{\mathrm{STF}}\left(M_i\mathcal{E}_{j\langle a} - S_i\mathcal{B}_{j\langle a}\right)\hat{\omega}_{b\rangle}{}^{ij} - \tfrac{8}{3}\epsilon^{pq}{}_{\langle j}\left(2\mathcal{E}_{k\rangle p}M_q + \mathcal{B}_{k\rangle p}S_q\right)\epsilon^k{}_{i(a}\hat{\omega}_{b)}{}^{ij} \\ &\quad + \tfrac{16}{15}m^2\mathcal{E}_{ab} + \tfrac{4}{15}\left(29M_{\langle a}\mathcal{E}_{bi\rangle} - 14S_{\langle a}\mathcal{B}_{bi\rangle}\right)\omega^i \\ &\quad - \tfrac{16}{45}\,\underset{ab}{\mathrm{STF}}\,\epsilon_{ai}{}^j\omega^i\epsilon^{pq}{}_{(b}\left(13\mathcal{E}_{j)q}M_p + 14\mathcal{B}_{j)q}S_p\right).\end{aligned} \tag{A.21}$$

B STF multipole decompositions

All formulas in this Appendix are either taken directly from Refs. [187] and [180] or are easily derivable from formulas therein.

Any Cartesian tensor field depending on two angles θ^A spanning a sphere can be expanded in a unique decomposition in symmetric trace-free tensors. Such a decomposition is equivalent to a decomposition in tensorial harmonics, but it is sometimes more convenient. It begins with the fact that the angular dependence of a Cartesian tensor $T_S(\theta^A)$ can be expanded in a series of the form

$$T_S(\theta^A) = \sum_{\ell\geq 0} T_{S\langle L\rangle}\hat{\omega}^L, \tag{B.1}$$

where S and L denote multi-indices $S = i_1 \dots i_s$ and $L = j_1 \dots j_\ell$, angular brackets denote an STF combination of indices, ω^a is a Cartesian unit vector, $\omega^L := \omega^{j_1} \dots \omega^{j_\ell}$, and $\hat{\omega}^L := \omega^{\langle L\rangle}$. This is entirely equivalent to an expansion in spherical harmonics. Each coefficient $T_{S\langle L\rangle}$ can be found from the formula

$$T_{S\langle L\rangle} = \frac{(2\ell+1)!!}{4\pi\ell!}\int T_S(\theta^A)\hat{\omega}_L d\Omega, \tag{B.2}$$

where the double factorial is defined by $x!! = x(x-2)\cdots 1$. These coefficients can then be decomposed into irreducible STF tensors. For example, for $s = 1$, we have

$$T_{a\langle L\rangle} = \hat{T}^{(+1)}_{aL} + \epsilon^j{}_{a\langle i_\ell}\hat{T}^{(0)}_{L-1\rangle j} + \delta_{a\langle i_\ell}\hat{T}^{(-1)}_{L-1\rangle}, \tag{B.3}$$

where the $\hat{T}^{(n)}$'s are STF tensors given by

$$\hat{T}^{(+1)}_{L+1} := T_{\langle L+1\rangle}, \tag{B.4}$$

$$\hat{T}^{(0)}_{L} := \frac{\ell}{\ell+1}T_{pq\langle L-1}\epsilon_{i_\ell\rangle}{}^{pq}, \tag{B.5}$$

$$\hat{T}^{(-1)}_{L-1} := \frac{2\ell-1}{2\ell+1}T^j{}_{jL-1}. \tag{B.6}$$

Similarly, for a symmetric tensor T_S with $s = 2$, we have

$$\begin{aligned}T_{ab\langle L\rangle} = \underset{L}{\mathrm{STF}}\,\underset{ab}{\mathrm{STF}}\Big(&\epsilon^p{}_{ai_\ell}\hat{T}^{(+1)}_{bpL-1} + \delta_{ai_\ell}\hat{T}^{(0)}_{bL-1} + \delta_{ai_\ell}\epsilon^p{}_{bi_{\ell-1}}\hat{T}^{(-1)}_{pL-2} + \delta_{ai_\ell}\delta_{bi_{\ell-1}}\hat{T}^{(-2)}_{L-2}\Big)\\ &+ \hat{T}^{(+2)}_{abL} + \delta_{ab}\hat{K}_L,\end{aligned} \tag{B.7}$$

where

$$\hat{T}^{(+2)}_{L+2} := T_{\langle L+2\rangle}, \tag{B.8}$$

$$\hat{T}^{(+1)}_{L+1} := \frac{2\ell}{\ell+2}\underset{L+1}{\mathrm{STF}}(T_{\langle pi_\ell\rangle qL-1}\epsilon_{i_{\ell+1}}{}^{pq}), \tag{B.9}$$

$$\hat{T}^{(0)}_{L} := \frac{6\ell(2\ell-1)}{(\ell+1)(2\ell+3)}\underset{L}{\mathrm{STF}}(T_{\langle ji_\ell\rangle}{}^j{}_{L-1}), \tag{B.10}$$

$$\hat{T}^{(-1)}_{L-1} := \frac{2(\ell-1)(2\ell-1)}{(\ell+1)(2\ell+1)}\underset{L-1}{\mathrm{STF}}(T_{\langle jp\rangle q}{}^j{}_{L-2}\epsilon_{i_{\ell-1}}{}^{pq}), \tag{B.11}$$

$$\hat{T}^{(-2)}_{L-2} := \frac{2\ell-3}{2\ell+1}T_{\langle jk\rangle}{}^{jk}{}_{L-2} \tag{B.12}$$

$$\hat{K}_L := \tfrac{1}{3}T^j{}_{jL}. \tag{B.13}$$

These decompositions are equivalent to the formulas for addition of angular momenta, $J = S + L$, which results in terms with angular momentum $\ell - s \leq j \leq \ell + s$; the superscript labels $(\pm n)$ in these formulas indicate by how much each term's angular momentum differs from ℓ.

By substituting Eqs. (B.3) and (B.7) into Eq. (B.1), we find that a scalar, a Cartesian 3-vector, and the symmetric part of a rank-2 Cartesian 3-tensor can be decomposed as, respectively,

$$T(\theta^A) = \sum_{\ell\geq 0} \hat{A}_L \hat{\omega}^L, \tag{B.14}$$

$$T_a(\theta^A) = \sum_{\ell\geq 0} \hat{B}_L \hat{\omega}_{aL} + \sum_{\ell\geq 1} \left[\hat{C}_{aL-1}\hat{\omega}^{L-1} + \epsilon^i{}_{aj}\hat{D}_{iL-1}\hat{\omega}^{jL-1}\right], \tag{B.15}$$

$$\begin{aligned} T_{(ab)}(\theta^A) &= \delta_{ab}\sum_{\ell\geq 0}\hat{K}_L\hat{\omega}^L + \sum_{\ell\geq 0}\hat{E}_L\hat{\omega}_{ab}{}^L + \sum_{\ell\geq 1}\left[\hat{F}_{L-1\langle a}\hat{\omega}_{b\rangle}{}^{L-1} + \epsilon^{ij}{}_{(a}\hat{\omega}_{b)i}{}^{L-1}\hat{G}_{jL-1}\right] \\ &\quad + \sum_{\ell\geq 2}\left[\hat{H}_{abL-2}\hat{\omega}^{L-2} + \epsilon^{ij}{}_{(a}\hat{I}_{b)jL-2}\hat{\omega}_i{}^{L-2}\right]. \end{aligned} \tag{B.16}$$

Each term in these decompositions is algebraically independent of all the other terms.

We can also reverse a decomposition to "peel off" a fixed index from an STF expression:

$$(\ell+1)\,\underset{iL}{\mathrm{STF}}\,T_{i\langle L\rangle} = T_{i\langle L\rangle} + \ell\,\underset{L}{\mathrm{STF}}\,T_{i_\ell\langle iL-1\rangle} - \frac{2\ell}{2\ell+1}\,\underset{L}{\mathrm{STF}}\,T^j{}_{\langle jL-1\rangle}\delta_{i_\ell i}. \tag{B.17}$$

In evaluating the action of the wave operator on a decomposed tensor, the following formulas are useful:

$$\omega^c\hat{\omega}^L = \hat{\omega}^{cL} + \frac{\ell}{2\ell+1}\delta^{c\langle i_1}\hat{\omega}^{i_2\ldots i_\ell\rangle}, \tag{B.18}$$

$$\omega_c\hat{\omega}^{cL} = \frac{\ell+1}{2\ell+1}\hat{\omega}^L, \tag{B.19}$$

$$r\partial_c\hat{\omega}_L = -\ell\hat{\omega}_{cL} + \frac{\ell(\ell+1)}{2\ell+1}\delta_{c\langle i_1}\hat{\omega}_{i_2\ldots i_\ell\rangle}, \tag{B.20}$$

$$\partial^c\partial_c\hat{\omega}^L = -\frac{\ell(\ell+1)}{r^2}\hat{\omega}^L, \tag{B.21}$$

$$\omega^c\partial_c\hat{\omega}^L = 0, \tag{B.22}$$

$$r\partial_c\hat{\omega}^{cL} = \frac{(\ell+1)(\ell+2)}{(2\ell+1)}\hat{\omega}^L. \tag{B.23}$$

In evaluating the t-component of the Lorenz gauge condition, the following formula is useful for finding the most divergent term (in an expansion in r):

$$r\partial^c h^{(n,m)}_{tc} = \sum_{\ell\geq 0}\frac{(\ell+1)(\ell+2)}{2\ell+1}\hat{B}^{(n,m)}_L\hat{\omega}^L - \sum_{\ell\geq 2}(\ell-1)\hat{C}^{(n,m)}_L\hat{\omega}^L. \tag{B.24}$$

And in evaluating the a-component, the following formula is useful for the same purpose:

$$\begin{aligned} &r\partial^b h^{(n,m)}_{ab} - \tfrac{1}{2}r\eta^{\beta\gamma}\partial_a h^{(n,m)}_{\beta\gamma} \\ &= \sum_{\ell\geq 0}\left[\tfrac{1}{2}\ell(\hat{K}^{(n,m)}_L - \hat{A}^{(n,m)}_L) + \frac{(\ell+2)(\ell+3)}{2\ell+3}\hat{E}^{(n,m)}_L - \tfrac{1}{6}\ell\hat{F}^{(n,m)}_L\right]\hat{\omega}_a{}^L \\ &\quad + \sum_{\ell\geq 1}\left[\frac{\ell(\ell+1)}{2(2\ell+1)}(\hat{A}^{(n,m)}_{aL-1} - \hat{K}^{(n,m)}_{aL-1}) + \frac{(\ell+1)^2(2\ell+3)}{6(2\ell+1)(2\ell-1)}\hat{F}^{(n,m)}_{aL-1}\right. \\ &\quad \left. - (\ell-2)\hat{H}^{(n,m)}_{aL-1}\right]\hat{\omega}^{L-1} + \sum_{\ell\geq 1}\left[\frac{(\ell+2)^2}{2(2\ell+1)}\hat{G}^{(n,m)}_{dL-1} - \tfrac{1}{2}(\ell-1)\hat{I}^{(n,m)}_{dL-1}\right]\epsilon_{ac}{}^d\hat{\omega}^{cL-1} \end{aligned} \tag{B.25}$$

where we have defined $\hat{H}^{(n,m)}_a := 0$ and $\hat{I}^{(n,m)}_a := 0$.

The unit vector ω_i satisfies the following integral identities:

$$\int \hat{\omega}_L d\Omega = 0 \text{ if } \ell > 0, \tag{B.26}$$

$$\int \omega_L d\Omega = 0 \text{ if } \ell \text{ is odd}, \tag{B.27}$$

$$\int \omega_L d\Omega = 4\pi \frac{\delta_{\{i_1 i_2} \dots \delta_{i_{\ell-1} i_\ell\}}}{(\ell+1)!!} \text{ if } \ell \text{ is even}, \tag{B.28}$$

where the curly braces indicate the smallest set of permutations of indices that make the result symmetric. For example, $\delta_{\{ab}\omega_{c\}} = \delta_{ab}\omega_c + \delta_{bc}\omega_a + \delta_{ca}\omega_b$.

References

[1] P. Dirac, *Classical theory of radiating electrons*, Proc. R. Soc. London, Ser. A **167**, 148 (1938).

[2] S. E. Gralla, A. I. Harte, and R. M. Wald, *A rigorous derivation of electromagnetic self-force*, Phys. Rev. **D80**, 024031 (2009).

[3] F. Rohrlich, *Classical charged particles* (Addison-Wesley, Redwood City, CA, U.S.A., 1990).

[4] B. DeWitt and R. Brehme, *Radiation Damping in a Gravitational Field*, Ann. Phys. (N.Y.) **9**, 220 (1960).

[5] J. Hobbs, *A vierbein formalism for radiation damping*, Ann. Phys. (N.Y.) **47**, 141 (1968).

[6] Y. Mino, M. Sasaki, and T. Tanaka, *Gravitational radiation reaction to a particle motion*, Phys. Rev. D **55**, 3457 (1997).

[7] T. Quinn and R. Wald, *Axiomatic approach to electromagnetic and gravitational radiation reaction of particles in curved spacetime*, Phys. Rev. D **56**, 3381 (1997).

[8] T. Quinn, *Axiomatic approach to radiation reaction of scalar point particles in curved spacetime*, Phys. Rev. D **62**, 064029 (2000).

[9] J. Jackson, *Classical Electrodynamics, Third Edition* (Wiley, New York, NY, U.S.A., 1999).

[10] C. Teitelboim, D. Villarroel, and C. van Weert, *Classical electrodynamics of retarded fields and point particles*, Riv. Nuovo Cimento **3**, 9 (1980).

[11] L. D. Landau and E. M. Lifshitz, *The classical theory of fields, fourth edition* (Butterworth-Heinemann, Oxford, England, 2000).

[12] E. Flanagan and R. Wald, *Does back reaction enforce the averaged null energy condition in semiclassical gravity?*, Phys. Rev. D **54**, 6233 (1996).

[13] S. Detweiler and B. Whiting, *Self-force via a Green's function decomposition*, Phys. Rev. D **67**, 024025 (2003).

[14] J. Synge, *Relativity: The general theory* (North-Holland, Amsterdam, The Netherlands, 1960).

[15] E. Poisson, *Retarded coordinates based at a world line and the motion of a small black hole in an external universe*, Phys. Rev. D **69**, 084007 (2004).

[16] S. E. Gralla and R. M. Wald, *A rigorous derivation of gravitational self-force*, Class. Quant. Grav. **25**, 205009 (2008).

[17] A. Pound, *Self-consistent gravitational self-force*, Phys. Rev. D **81**, 024023 (2010).

[18] A. Pound, *Singular perturbation techniques in the gravitational self-force problem*, Phys. Rev. D **81**, 124009 (2010).

[19] L. Barack and A. Ori, *Gravitational self-force and gauge transformations*, Phys. Rev. D **64**, 124003 (2001).

[20] LISA Home Page: http://lisa.jpl.nasa.gov.

[21] A. Smith and C. Will, *Force on a static charge outside a Schwarzschild black hole*, Phys. Rev. D **22**, 1276 (1980).

[22] A. I. Zel'nikov and V. P. Frolov, *Influence of gravitation on the self-energy of charged particles*, Sov. Phys. JETP **55**, 919 (1982).

[23] E. Copson, *On electrostatics in a gravitational field*, Proc. R. Soc. London, Ser. A **116**, 720 (1928).

[24] B. Linet, *Electrostatics and magnetostatics in Schwarzschild metric*, J. Phys. A **9**, 1081 (1976).

[25] B. Leaute and B. Linet, *Electrostatics in a Reissner-Nordström space-time*, Phys. Lett. A **58**, 5 (1976).

[26] A. Wiseman, *Self-force on a static scalar test charge outside a Schwarzschild black hole*, Phys. Rev. D **61**, 084014 (2000).

[27] C. Misner, K. Thorne, and J. Wheeler, *Gravitation* (Freeman, San Francisco, CA, U.S.A., 1973).

[28] F. Manasse and C. Misner, *Fermi normal coordinates and some basic concepts in differential geometry*, J. Math. Phys. **4**, 735 (1963).

[29] J. Hadamard, *Lectures on Cauchy's problem in linear partial differential equations* (Yale University Press, New Haven, CT, U.S.A., 1923).

[30] F. Friedlander, *The wave equation on a curved spacetime* (Cambridge University Press, Cambridge, U.K., 1975).

[31] L. Barack, *Gravitational self force in extreme mass-ratio inspirals*, Class. Quantum Grav. **26**, 213001 (2009).

[32] C. Lousto (editor), *Gravitational radiation from binary black holes: advances in the perturbative approach*, volume 22 (Class. Quantum Grav., 2005).

[33] S. Detweiler, *Perspective on gravitational self-force analyses*, Class. Quantum Grav. **22**, 681 (2005).

[34] B. Whiting, *Identifying the singular field for self-force evaluation*, Class. Quantum Grav. **22**, S661 (2005).

[35] C. Morette-DeWitt and B. DeWitt, *Falling charges*, Physics (Long Island City, N.Y.) **1**, 3 (1964).

[36] M. Pfenning and E. Poisson, *Scalar, electromagnetic, and gravitational self-forces in weakly curved spacetimes*, Phys. Rev. D **65**, 084001 (2002).

[37] L. Barack and A. Ori, *Mode sum regularization approach for the self-force in black hole spacetime*, Phys. Rev. D **61**, 061502 (2000).

[38] L. Barack, *Self-force on a scalar particle in spherically symmetric spacetime via mode-sum regularization: Radial trajectories*, Phys. Rev. D **62**, 084027 (2000).

[39] L. Barack, Y. Mino, H. Nakano, A. Ori, and M. Sasaki, *Calculating the gravitational self-Force in Schwarzschild spacetime*, Phys. Rev. Lett. **88**, 091101 (2002).

[40] L. Barack, *Gravitational self-force by mode sum regularization*, Phys. Rev. D **64**, 084021 (2001).

[41] L. Barack and A. Ori, *Regularization parameters for the self-force in Schwarzschild spacetime: Scalar case*, Phys. Rev. D **66**, 084022 (2002).

[42] L. Barack and A. Ori, *Regularization parameters for the self-force in Schwarzschild spacetime. II. Gravitational and electromagnetic cases*, Phys. Rev. D **67**, 024029 (2003).

[43] L. Barack and A. Ori, *Gravitational self-force on a particle orbiting a Kerr black hole*, Phys. Rev. Lett. **90**, 111101 (2003).

[44] Y. Mino, H. Nakano, and M. Sasaki, *Covariant self-force regularization of a particle orbiting a Schwarzschild black hole – Mode decomposition regularization*, Prog. Theor. Phys. **108**, 1039 (2003).

[45] C. Lousto, *Pragmatic approach to gravitational radiation reaction in binary black holes*, Phys. Rev. Lett. **84**, 5251 (2000).

[46] S. Detweiler, E. Messaritaki, and B. Whiting, *Self-force of a scalar field for circular orbits about a Schwarzschild black hole*, Phys. Rev. D **67**, 104016 (2003).

[47] L. Burko, A. Harte, and E. Poisson, *Mass loss by a scalar charge in an expanding universe*, Phys. Rev. D **65**, 124006 (2002).

[48] R. Haas and E. Poisson, *Mass change and motion of a scalar charge in cosmological spacetimes*, Class. Quantum Grav. **22**, S739 (2005).

[49] K. Shankar and B. Whiting, *Self-force of a static electric charge near a Schwarzschild star*, Phys. Rev. D **76**, 124027 (2007).

[50] R. Haas and E. Poisson, *Mode-sum regularization of the scalar self-force: Formulation in terms of a tetrad decomposition of the singular field*, Phys. Rev. D **74**, 004009 (2006).

[51] L. Burko, *Self-force on static charges in Schwarzschild spacetime*, Class. Quantum Grav. **17**, 227 (2000).

[52] L. Burko, Y. Liu, and Y. Soen, *Self-force on charges in the spacetime of spherical shells*, Phys. Rev. D **63**, 024015 (2001).

[53] L. Barack and L. Burko, *Radiation-reaction force on a particle plunging into a black hole*, Phys. Rev. D **62**, 084040 (2000).

[54] L. Burko, *Self-force on a particle in orbit around a black hole*, Phys. Rev. Lett. **84**, 4529 (2000).

[55] P. Canizares and C. Sopuerta, *Efficient pseudospectral method for the computation of the self-force on a charged particle: Circular geodesics around a Schwarzschild black hole*, Phys. Rev. D **79**, 084020 (2009).

[56] R. Haas, *Scalar self-force on eccentric geodesics in Schwarzschild spacetime: a time-domain computation*, Phys. Rev. D **75**, 124011 (2007).

[57] R. Haas, *Self-force on point particles in orbit around a schwarzschild black hole*, Ph.D. thesis, University of Guelph (2008).

[58] L. Barack and N. Sago, *Gravitational self-force on a particle in circular orbit around a Schwarzschild black hole*, Phys. Rev. D **75**, 064021 (2007).

[59] S. Akcay, *Fast frequency-domain algorithm for gravitational self-force: Circular orbits in Schwarzschild spacetime*, Phys. Rev. D **83**, 124026 (2011).

[60] L. Barack and N. Sago, *Gravitational self-force correction to the innermost stable circular orbit of a Schwarzschild black hole*, Phys. Rev. Lett. **102**, 191101 (2009).

[61] S. Detweiler, *Consequence of the gravitational self-force for circular orbits of the Schwarzschild geometry*, Phys. Rev. D **77**, 124026 (2008).

[62] N. Sago, L. Barack, and S. Detweiler, *Two approaches for the gravitational self force in black hole spacetime: Comparison of numerical results*, Phys. Rev. D **78**, 124024 (2008).

[63] L. Barack, T. Damour, and N. Sago, *Precession effect of the gravitational self-force in a Schwarzschild spacetime and the effective one-body formalism*, Phys. Rev. D **82**, 084036 (2010).

[64] L. Blanchet, S. Detweiler, A. Le Tiec, and B. Whiting, *Post-Newtonian and numerical calculations of the gravitational self-force for circular orbits in the Schwarzschild geometry*, Phys. Rev. D **81**, 064004 (2010).

[65] L. Blanchet, S. Detweiler, A. Le Tiec, and B. Whiting, *High-order post-Newtonian fit of the gravitational self-force for circular orbits in the Schwarzschild geometry*, Phys. Rev. D **81**, 084033 (2010).

[66] T. Damour, *Gravitational self-force in a Schwarzschild background and the effective one-body formalism*, Phys. Rev. D **81**, 024017 (2010).

[67] L. Burko and Y. Liu, *Self-force on a scalar charge in the spacetime of a stationary, axisymmetric black hole*, Phys. Rev. D **64**, 024006 (2001).

[68] N. Warburton and L. Barack, *Self-force on a scalar charge in Kerr spacetime: Circular equatorial orbits*, Phys. Rev. D **81**, 084039 (2010).

[69] S. Teukolsky, *Perturbations of a rotating black hole. I. Fundamental equations for gravitational, electromagnetic, and neutrino-field perturbations*, Astrophys. J. **185**, 635 (1973).

[70] P. Chrzanowski, *Vector potential and metric perturbations of a rotating black hole*, Phys. Rev. D **11**, 2042 (1975).

[71] J. Cohen and L. Kegeles, *Electromagnetic fields in curved spaces: a constructive procedure*, Phys. Rev. D **10**, 1070 (1974).

[72] L. Kegeles and J. Cohen, *Constructive procedure for perturbations of spacetimes*, Phys. Rev. D **19**, 1641 (1979).

[73] J. Stewart, *Hertz-Bromwich-Debye-Whittaker-Penrose potentials in general relativity*, Proc. Roy. Soc. London, Ser. A **367**, 527 (1979).

[74] R. Wald, *Construction of solutions of gravitational, electromagnetic, or other perturbation equations from solutions of decoupled equations*, Phys. Rev. Lett. **41**, 203 (1978).

[75] C. Lousto and B. Whiting, *Reconstruction of black hole metric perturbations from the Weyl curvature*, Phys. Rev. D **66**, 024026 (2002).

[76] A. Ori, *Reconstruction of inhomogeneous metric perturbations and electromagnetic four-potential in Kerr spacetime*, Phys. Rev. D **67**, 124010 (2003).

[77] B. Whiting and L. Price, *Metric reconstruction from Weyl scalars*, Class. Quantum Grav. **22**, S589 (2005).

[78] T. Keidl, J. Friedman, and A. Wiseman, *Finding fields and self-force in a gauge appropriate to separable wave equations*, Phys. Rev. D **75**, 124009 (2006).

[79] T. Keidl, *Self-force for extreme mass ratio inspiral*, Ph.D. thesis, University of Wisconsin-Milwaukee (2008).

[80] T. Keidl, A. Shah, J. Friedman, D. Kim, and L. Price, *Gravitational self-force in a radiation gauge*, Phys. Rev. D **82**, 124012 (2010).

[81] S. Detweiler and E. Poisson, *Low multipole contributions to the gravitational self-force*, Phys. Rev. D **69**, 084019 (2004).

[82] J. Barton, D. Lazar, D. Kennefick, G. Khanna, and L. Burko, *Computational efficiency of frequency– and time–domain calculations of extreme mass–ratio binaries: Equatorial orbits*, Phys. Rev. D **78**, 064042 (2010).

[83] R. Lopez-Aleman, G. Khanna, and J. Pullin, *Perturbative evolution of particle orbits around Kerr black holes: time domain calculation*, Class. Quantum Grav. **20**, 3259 (2003).

[84] G. Khanna, *Teukolsky evolution of particle orbits around Kerr black holes in the time domain: Elliptic and inclined orbits*, Phys. Rev. D **69**, 024016 (2006).

[85] L. Burko and G. Khanna, *Accurate time-domain gravitational waveforms for extreme-mass-ratio binaries*, Europhys. Lett. **78**, 60005 (2007).

[86] P. Sundararajan, G. Khanna, and S. Hughes, *Towards adiabatic waveforms for inspiral into Kerr black holes: I. A new model of the source for the time domain perturbation equation*, Phys. Rev. D **76**, 104005 (2007).

[87] C. Lousto and R. Price, *Understanding initial data for black hole collisions*, Phys. Rev. D **56**, 6439 (1997).

[88] K. Martel and E. Poisson, *One-parameter family of time-symmetric initial data for the radial infall of a particle into a Schwarzschild black hole*, Phys. Rev. D **66**, 084001 (2002).

[89] I. Vega and S. Detweiler, *Regularization of fields for self-force problems in curved spacetime: Foundations and a time-domain application*, Phys. Rev. D **77**, 084008 (2008).

[90] S. Field, J. Hesthaven, and S. Lau, *Discontinuous Galerkin method for computing gravitational waveforms from extreme mass ratio binaries*, Class. Quantum Grav. **26**, 165010 (2009).

[91] S. Field, J. Hesthaven, and S. Lau, *Persistent junk solutions in time-domain modeling of extreme mass ratio binaries*, Phys. Rev. D **81**, 124030 (2010).

[92] P. Canizares, C. Sopuerta, and J. Jaramillo, *Pseudospectral collocation methods for the computation of the self-force on a charged particle: Generic orbits around a Schwarzschild black hole*, Phys. Rev. D **82**, 044023 (2010).

[93] L. Barack and D. Golbourn, *Scalar-field perturbations from a particle orbiting a black hole using numerical evolution in 2+1 dimensions*, Phys. Rev. D **76**, 044020 (2007).

[94] L. Barack, D. Golbourn, and N. Sago, *m-mode regularization scheme for the self-force in Kerr spacetime*, Phys. Rev. D **76**, 124036 (2007).

[95] S. Dolan and L. Barack, *Self-force via m-mode regularization and 2+1D evolution: Foundations and a scalar-field implementation on Schwarzschild spacetime*, Phys. Rev. D **83**, 124019 (2011).

[96] I. Vega, P. Diener, W. Tichy, and S. Detweiler, *Self-force with (3+1) codes: a primer for numerical relativists*, Phys. Rev. D **80**, 084021 (2009).

[97] I. Vega, *The dynamics of point particles around black holes*, Ph.D. thesis, University of Florida (2009).

[98] C. Lousto and H. Nakano, *A new method to integrate (2+1)-wave equations with Dirac's delta functions as sources*, Class. Quantum Grav. **25**, 145018 (2008).

[99] P. Anderson and B. Hu, *Radiation reaction in Schwarzschild spacetime: Retarded Greens function via Hadamard-WKB expansion*, Phys. Rev. D **69**, 064039 (2004).

[100] P. Anderson, A. Eftekharzadeh, and B. Hu, *Self-force on a scalar charge in radial infall from rest using the Hadamard-WKB expansion*, Phys. Rev. D **73**, 064023 (2006).

[101] W. Anderson and A. Wiseman, *A matched expansion approach to practical self-force calculations*, Class. Quantum Grav. **22**, S783 (2005).

[102] W. Anderson, E. Flanagan, and A. Ottewill, *Quasilocal contribution to the gravitational self-force*, Phys. Rev. D **71**, 024036 (2005).

[103] A. Ottewill and B. Wardell, *Quasilocal contribution to the scalar self-force: Geodesic motion*, Phys. Rev. D **77**, 104002 (2008).

[104] B. Wardell, *Green functions and radiation reaction from a spacetime perspective*, Ph.D. thesis, University College Dublin (2009).

[105] A. Ottewill and B. Wardell, *Quasilocal contribution to the scalar self-force: Nongeodesic motion*, Phys. Rev. D **79**, 024031 (2009).

[106] M. Casals, S. Dolan, A. Ottewill, and B. Wardell, *Padé approximants of the Green function in spherically symmetric spacetimes*, Phys. Rev. D **79**, 124044 (2009).

[107] Y. Decanini and A. Folacci, *Off-diagonal coefficients of the DeWitt-Schwinger and Hadamard representations of the Feynman propagator*, Phys. Rev. D **73**, 044027 (2006).

[108] M. Casals, S. Dolan, A. Ottewill, and B. Wardell, *Self-force calculations with matched expansions and quasinormal mode sums*, Phys. Rev. D **79**, 124043 (2009).

[109] P. Peters and J. Mathews, *Gravitational radiation from point masses in a Keplerian orbit*, Phys. Rev. **131**, 435 (1963).

[110] P. Peters, *Gravitational radiation and the motion of two point masses*, Phys. Rev. **136**, B1224 (1964).

[111] Y. Mino, *Perturbative approach to an orbital evolution around a supermassive black hole*, Phys. Rev. D **67**, 084027 (2003).

[112] Y. Mino, *Self-force in the radiation reaction formula – adiabatic approximation of a metric perturbation and an orbit*, Prog. Theor. Phys. **113**, 733 (2005).

[113] Y. Mino, *Adiabatic expansion for a metric perturbation and the condition to solve the gauge problem for gravitational radiation reaction problem*, Prog. Theor. Phys. **115**, 43 (2006).

[114] S. Hughes, *Evolution of circular, nonequatorial orbits of Kerr black holes due to gravitational-wave emission*, Phys. Rev. D **61**, 084004 (2000).

[115] S. Drasco, E. Flanagan, and S. Hughes, *Computing inspirals in Kerr in the adiabatic regime: I. The scalar case*, Class. Quantum Grav. **22**, S801 (2005).

[116] S. Drasco and S. Hughes, *Gravitational wave snapshots of generic extreme mass ratio inspirals*, Phys. Rev. D **73**, 024027 (2006).

[117] N. Sago, T. Tanaka, W. Hikida, and H. Nakano, *Adiabatic radiation reaction to orbits in Kerr spacetime*, Prog. Theor. Phys. **114**, 509 (2005).

[118] N. Sago, T. Tanaka, W. Hikida, K. Ganz, and H. Nakano, *Adiabatic evolution of orbital parameters in the Kerr spacetime*, Prog. Theor. Phys. **115**, 873 (2006).

[119] K. Ganz, W. Hikida, H. Nakano, N. Sago, and T. Tanaka, *Adiabatic evolution of three 'constants' of motion for greatly inclined orbits in Kerr spacetime*, Prog. Theor. Phys. **117**, 1041 (2007).

[120] Y. Mino and R. Price, *Two-timescale adiabatic expansion of a scalar field model*, Phys. Rev. D **77**, 064001 (2008).

[121] T. Hinderer and E. E. Flanagan, *Two timescale analysis of extreme mass ratio inspirals in Kerr. I. Orbital Motion*, Phys. Rev. D **78**, 064028 (2008).

[122] D. Gal'tsov, *Radiation reaction in the Kerr gravitational field*, J. Phys. A: Math. Gen. **15**, 3737 (1982).

[123] S. Hughes, S. Drasco, E. Flanagan, and J. Franklin, *Gravitational radiation reaction and inspiral waveforms in the adiabatic limit*, Phys. Rev. Lett. **94**, 221101 (2005).

[124] L. M. Burko, *The importance of conservative self forces for binaries undergoing radiation damping*, Int. J. Mod. Phys. A **16**, 1471 (2001).

[125] A. Pound, E. Poisson, and B. Nickel, *Limitations of the adiabatic approximation to the gravitational self-force*, Phys. Rev. D **72**, 124001 (2005).

[126] A. Pound and E. Poisson, *Multi-scale analysis of the electromagnetic self-force in a weak gravitational field*, Phys. Rev. D **77**, 044012 (2008).

[127] A. Pound and E. Poisson, *Osculating orbits in Schwarzschild spacetime, with an application to extreme mass-ratio inspirals*, Phys. Rev. D **77**, 044013 (2008).

[128] E. A. Huerta and J. R. Gair, *Influence of conservative corrections on parameter estimation for EMRIs*, Phys. Rev. D **79**, 084021 (2009).

[129] A. Pound, *Motion of small bodies in general relativity: foundations and implementations of the self-force*, Ph.D. thesis, University of Guelph (2010).

[130] E. E. Flanagan and T. Hinderer, *Transient resonances in the inspirals of point particles into black holes*, Online Los Alamos Archive Preprint, http://arxiv.org/abs/1009.4923 (2010).

[131] J. Kevorkian and J. D. Cole, *Multiple scale and singular perturbation methods* (Springer, New York, 1996).

[132] L. Barack and C. Lousto, *Computing the gravitational self-force on a compact object plunging into a Schwarzschild black hole*, Phys. Rev. D **66**, 061502 (2002).

[133] L. Barack and N. Sago, *Gravitational self-force on a particle in eccentric orbit around a Schwarzschild black hole*, Phys. Rev. D **81**, 084021 (2010).

[134] A. Ori and K. Thorne, *Transition from inspiral to plunge for a compact body in a circular equatorial orbit around a massive, spinning black hole*, Phys. Rev. D **62**, 124022 (2000).

[135] P. Sundararajan, *Transition from adiabatic inspiral to geodesic plunge for a compact object around a massive Kerr black hole: Generic orbits*, Phys. Rev. D **77**, 124050 (2008).

[136] L. Diaz-Rivera, E. Messaritaki, B. Whiting, and S. Detweiler, *Scalar field self-force effects on orbits about a Schwarzschild black hole*, Phys. Rev. D **70**, 124018 (2004).

[137] L. Barack and N. Sago, *Beyond the geodesic approximation: Conservative effects of the gravitational self-force in eccentric orbits around a Schwarzschild black hole*, Phys. Rev. D **83**, 084023 (2011).

[138] K. Thorne and J. Hartle, *Laws of motion and precession for black holes and other bodies*, Phys. Rev. D **31**, 1815 (1985).

[139] X.-H. Zhang, *Multipole expansions of the general-relativistic gravitational field of the external universe*, Phys. Rev. D **34**, 991 (1986).

[140] D. Sciama, P. Waylen, and R. Gilman, *Generally covariant integral formulation of Einstein's field equations*, Phys. Rev. **187**, 1762 (1969).

[141] T. Quinn and R. Wald, *Energy conservation for point particles undergoing radiation reaction*, Phys. Rev. D **60**, 064009 (1999).

[142] R. Geroch and J. Traschen, *Strings and other distributional sources in general relativity*, Phys. Rev. D **36**, 1017 (1987).

[143] M. Walker and C. M. Will, *The approximation of radiative effects in relativistic gravity — Gravitational radiation reaction and energy loss in nearly Newtonian systems*, Astrophys. J. **242**, L129 (1980).

[144] L. Blanchet, *Gravitational radiation from post-Newtonian sources and inspiralling compact binaries*, Living Rev. Relativity **9**(4) (2006).

[145] W. Eckhaus, *Asymptotic analysis of singular perturbations* (Elsevier North-Holland, New York, 1979).

[146] M. H. Holmes, *Introduction to perturbation methods* (Springer-Verlag, New York, 1995).

[147] P. A. Lagerstrom, *Matched asymptotic expansions* (Springer-Verlag, New York, 1988).

[148] F. Verhulst, *Methods and applications of singular perturbations* (Springer, New York, 2005).

[149] P. Havas, *Radiation damping in general relativity*, Phys. Rev. **108**(5), 1351 (1957).

[150] P. Havas and J. N. Goldberg, *Lorentz-invariant equations of motion of point masses in the general theory of relativity*, Phys. Rev. **128**(1), 398 (1962).

[151] C. Morette-DeWitt and J. Ging, *Freinage dû à la radiation gravitationnelle*, C. R. Hebd. Seanc. Acad. Sci. **251**, 1868 (1960).

[152] R. Steinbauer and J. A. Vickers, *The use of generalised functions and distributions in general relativity*, Class. Quant. Grav. **23**, R91 (2006).

[153] L. Infeld and A. Schild, *On the motion of test particles in general relativity*, Rev. Mod. Phys. **21**, 408 (1949).

[154] R. Geroch and P. Jang, *Motion of a body in general relativity*, J. Math. Phys. **16**, 65 (1975).

[155] J. Ehlers and R. Geroch, *Equation of motion of small bodies in relativity*, Ann. Phys. (N.Y.) **309**, 232 (2004).

[156] Y. Itoh and T. Futamase, *The post-Newtonian approximation for relativistic compact binaries*, Living Reviews in Relativity **10**(2) (2007).

[157] A. I. Harte, *Self-forces from generalized Killing fields*, Class. Quant. Grav. **25**, 235020 (2008).

[158] A. I. Harte, *Electromagnetic self-forces and generalized Killing fields*, Class. Quant. Grav. **26**, 155015 (2009).

[159] W. G. Dixon, *Dynamics of extended bodies in general relativity. I. Momentum and angular momentum*, Proc. Roy. Soc. Lond. **A314**, 499 (1970).

[160] W. G. Dixon, *Dynamics of extended bodies in general relativity. II. Moments of the charge-current vector*, Proc. Roy. Soc. Lond. **A319**, 509 (1970).

[161] W. G. Dixon, *Dynamics of Extended Bodies in General Relativity. III. Equations of Motion*, Phil. Trans. R. Soc. Lond. A **277**, 59 (1974).

[162] C. R. Galley, B. L. Hu, and S.-Y. Lin, *Electromagnetic and gravitational self-force on a relativistic particle from quantum fields in curved space*, Phys. Rev. D **74**, 024017 (2006).

[163] C. R. Galley and B. L. Hu, *Self-force on extreme mass ratio inspirals via curved spacetime effective field theory*, Phys. Rev. D **79**, 064002 (2009).

[164] P. D'Eath, *Dynamics of a small black hole in a background universe*, Phys. Rev. D **11**, 1387 (1975).

[165] P. D'Eath, *Black holes: Gravitational interactions* (Clarendon Press, Oxford, U.K., 1996).

[166] R. Kates, *Motion of a small body through an external field in general relativity calculated by matched asymptotic expansions*, Phys. Rev. D **22**, 1853 (1980).

[167] R. E. Kates, *Motion of an electrically or magnetically charged body with possibly strong internal gravity through external electromagnetic and gravitational fields*, Phys. Rev. D **22**, 1879 (1980).

[168] Y. Mino, M. Sasaki, and T. Tanaka, *Gravitational radiation reaction*, Prog. Theor. Phys. Suppl. **128**, 373 (1997).

[169] K. Alvi, *Approximate binary-black-hole metric*, Phys. Rev. D **61**, 124013 (2000).

[170] S. Taylor and E. Poisson, *Nonrotating black hole in a post-Newtonian tidal environment*, Phys. Rev. D **78**, 084016 (2008).

[171] T. Futamase, P. A. Hogan, and Y. Itoh, *Equations of motion in general relativity of a small charged black hole*, Phys. Rev. **D78**, 104014 (2008).

[172] R. Kates, *Underlying structure of singular perturbations on manifolds*, Ann. Phys. (N.Y.) **132**, 1 (1981).

[173] A. Einstein and L. Infeld, *On the motion of particles in general relativity theory*, Can. J. Math. **1**, 209 (1949).

[174] E. Racine and E. E. Flanagan, *Post-1-Newtonian equations of motion for systems of arbitrarily structured bodies*, Phys. Rev. D **71**, 044010 (2005).

[175] K. S. Thorne, *Multipole expansions of gravitational radiation*, Rev. Mod. Phys. **52**, 299 (1980).

[176] R. Schattner, *The center of mass in general relativity*, Gen. Rel. Grav. **10**, 377 (1979).

[177] J. Ehlers and E. Rudolph, *Dynamics of extended bodies in general relativity center-of-mass description and quasirigidity*, Gen. Rel. Grav. **8**, 197 (1977).

[178] G. F. Roach, *Green's Functions* (Cambridge University Press, Cambridge, 1982).

[179] A. Papapetrou, *Spinning test-particles in general relativity. I*, Proc. Roy. Soc. London A **209**, 248 (1951).

[180] T. Damour and B. Iyer, *Multipole analysis for electromagnetism and linearized gravity with irreducible Cartesian tensors*, Phys. Rev. D **43**, 3259 (1991).

[181] A. I. Harte, *Effective stress-energy tensors, self-force, and broken symmetry*, Class. Quant. Grav. **27**, 135002 (2010).

[182] S. Gralla, *Comments on first and second order gravitational self-force* (2009), presentation at the 12th Capra Meeting on Radiation Reaction.

[183] T. Fukumoto, T. Futamase, and Y. Itoh, *On the equation of motion for a fast moving small object using the strong field point particle limit*, Prog. Theor. Phys. **116**, 423 (2006).

[184] E. Rosenthal, *Construction of the second-order gravitational perturbations produced by a compact object*, Phys. Rev. D **73**, 044034 (2006).

[185] E. Rosenthal, *Second-order gravitational self-force*, Phys. Rev. D **74**, 084018 (2006).

[186] J. Steinhoff and D. Puetzfeld, *Multipolar equations of motion for extended test bodies in General Relativity*, Phys. Rev. **D81**, 044019 (2010).

[187] T. Damour and L. Blanchet, *Radiative gravitational fields I.*, Phil. Trans. R. Soc. Lond. A **320**, 379 (1986).